MPRA
Munich Personal RePEc Archive

Short-term forecasting of the US unemployment rate

Maas, Benedikt

University of Hamburg

16 April 2019

Online at https://mpra.ub.uni-muenchen.de/94066/
MPRA Paper No. 94066, posted 23 May 2019 09:27 UTC

Short-term forecasting of the US unemployment rate

Benedikt Maas*

April 16, 2019

Abstract

This paper aims to assess whether Google search data is useful when predicting the US unemployment rate among other more traditional predictor variables. A weekly Google index is derived from the keyword "unemployment" and is used in diffusion index variants along with the weekly number of initial claims and monthly estimated latent factors. The unemployment rate forecasts are generated using MIDAS regression models that take into account the actual frequencies of the predictor variables. The forecasts are made in real-time and the forecasts of the best forecasting models exceed, for the most part, the root mean squared forecast error of two benchmarks. However, as the forecasting horizon increases, the forecasting performance of the best diffusion index variants decreases over time, which suggests that the forecasting methods proposed in this paper are most useful in the short-term.

JEL classification: C32, C53, C55, E32

*University of Hamburg, Department of Economics, Von-Melle-Park 5, 20146 Hamburg, Germany; E-mail: *Benedikt.Maas@uni-hamburg.de*

1. Introduction

In general, traditional labor statistics are available with at least a one-month lag. However, a more timely estimate of the unemployment rate is desirable for investors and policymakers, especially in times of economic uncertainty. An accurate prediction of the US unemployment rate has become even more important after the 2008/09 recession, especially since the Federal Reserve announced in December 2012 a shift of its monetary policy to a specific unemployment rate threshold. The so-called “Evans Rule” stated that “the Committee decided to keep the target range for the federal funds rate at 0 to 1/4 percent and currently anticipates that this exceptionally low range for the federal funds rate will be appropriate at least as long as the unemployment rate remains above 6-1/2 percent.”[1]

This paper investigates whether or not the information given in Google searches is useful to predict the US unemployment rate. The idea behind using search engine data is that if an increase in searches is observed in connection with unemployment, then this could this give an early indication of an increasing unemployment rate. The potential predictive power of Google search is used alongside other more traditional predictors. One of these is the number initial claims (IC). The IC is widely used in the literature as a predictor variable in unemployment rate forecasts.[2] The current state of the economy is also considered as a predictor in the forecasts. The state of the economy has a major impact on the unemployment rate: during a recession, an increase in the unemployment rate is expected; while during an upswing and prosperity phase, a decrease in the unemployment rate is expected. To take into account the current state of the economy, unobserved latent factors are derived from a macroeconomic database by principal components, as suggested in Stock and Watson (2002). These factors are intended to establish a link between the economic situation and the unemployment rate in the forecasting exercise in this paper.

Given that the Google and IC data are available on a weekly frequency, this paper uses weekly data to forecast the monthly US unemployment rate with three diffusion index (DI) variants after Stock and Watson (2002) based on the mixed data sampling (MIDAS) regression model introduced by Ghysels et al. (2006, 2007). In addition, factor augmented versions, where the monthly unobserved latent factors and the weekly data are combined, are also estimated. In general, the MIDAS framework allows us to combine variables of mixed frequencies in a regression model. Related studies have used monthly averages of the Google and IC data (D’Amuri and Marcucci, 2017), but—as empirically shown in Smith (2016) who applies the MIDAS approach

[1]For the official statement of the Federal Reserve’s Open Market Committee see https://www.federalreserve.gov/newsevents/pressreleases/monetary20121212a.htm.

[2]See for example Montgomery et al. (1998).

to forecast the unemployment rate in the UK—there is no need to adjust frequency to the target variable and thus lose valuable information. As stated generally in Andreou et al. (2010), there is no reason to ignore the fact that variables involved in empirical models are generated from processes of mixed frequencies and are used to estimate econometric models based on an aggregation scheme of equal weights because an equal weighting scheme can lead to information losses and thus to inefficient or biased estimates.

The forecasts in this paper are conducted in real-time and almost all exceed an autoregressive benchmark for each forecast horizon. However, the results show a mixed picture, in which a combination of predictor variables is most favorable because the best empirical results change from horizon to horizon. Comparing the MIDAS short-term forecasts with the forecasts of D'Amuri and Marcucci (2017), which are based on monthly averages of an alternative Google index, the models presented here obtain a lower root mean squared forecast error (RMSFE) for the shortest forecast horizons compared to this benchmark.

The rest of this paper is organised as follows. Section 2 gives a compact overview of the related literature dealing with the use of Internet data to forecast economic variables. Moreover, it focuses on potential pitfalls when using internet data and the choice of the keyword to obtain the Google index. Section 3 explains the econometric framework. Section 4 describes the data, the forecasting models, and the simulated real-time forecasting design. Section 5 states the empirical results, while Section 6 concludes.

2. The use of Internet search data in forecasting

2.1. Related literature

Internet search data has been used in a number of different research topics. In economics, Choi and Varian (2012) show that Google Trends data can help to forecast near-term values of economic indicators, such as automobile sales, travel destinations, consumer confidence and initial claims for unemployment benefits. Their paper inspired many economists to use Google Trends data to predict a variable that can be linked to the behavior of households. For example Vosen and Schmidt (2011) forecast consumption of goods, whereas Bangwayo-Skeete and Skeete (2015) and Yang et al. (2015) use Google data to predict future tourism demand. Wu and Brynjolfsson (2015) predict US housing prices and sales. Using a Markov-switching framework, Chen et al. (2015) use Google search data to improve the timeliness of business cycle turning point identification and they successfully nowcast the peak date within a month that the turning point occurred. In their analysis, they use the three keywords "recession", "foreclosure help", and "layoff", which represent the aggregated economy, the credit market, and the labor market, respectively. Liu et al. (2018) also use Internet search behavior to forecast Chinese GDP. However, because Google is not prevalent in China, the authors use data from its Chinese counterpart Baidu.

Considering inflation expectations, Guzmán (2011) proposes a real-time measure using search queries obtained from Google. She demonstrates that higher frequency measures tend to outperform standard lower frequency measures such as the SPF in tests of accuracy, predictive power and out-of-sample forecasts.

Dergiades et al. (2014) analyze whether Google and social media data influence European financial markets. They find that the data provide significant short-run information for the Greek-German and Irish-German government bond yield differential.

Considering the unemployment rate, McLaren and Shanbhogue (2011) analyze the labor market in the United Kingdom and compare standard autoregressive (AR) models to those augmented with Internet data, finding that the augmented models outperform the autoregressive benchmarks. Askitas and Zimmermann (2009) demonstrate strong correlations between Google keyword searches and unemployment rates for Germany, and Fondeur and Karamé (2013) find that including Google data improves youth unemployment predictions in France. Vicente et al. (2015) investigate the unemployment rate in Spain using autoregressive integrated moving average (ARIMA) models with an included explanatory variable that is derived from the Google search term "job offers". They find that significant forecasting improvements are observed when Google Trends variables are also included.

D'Amuri and Marcucci (2017) predict the US unemployment rate by obtaining a Google Trends index (GI) from the keyword "jobs", which is then used as an additional regressor in an autoregressive model. They consider the weekly availability of Google Trends data and use only the data from one specific week and the monthly averages in their forecasting models. Additionally, the Google data is aligned with the relevant weeks for the unemployment survey of the Bureau of Labor Statistics (BLS), to ensure that both variables capture the same information. Considering this data adjustment, they show that the Google-based forecast models outperform most of the considered competitor models.

Smith (2016) uses Google data to forecast the unemployment rate in the United Kingdom, where the author uses the keyword "redundancy" to obtain the monthly Google index. Using AR models with an additional explanatory variable, the forecasting performance of the Google index is compared to other competitor models. Two of the competing indicators are variables that are derived from small-scale factor models. The derived factors are created from the first principal component from a static principal component analysis (PCA). Additionally, the author uses weekly Google data as a high-frequent predictor in an unrestricted MIDAS regression model. Overall, it is found the data from Google Trends offer similar and at times better forecasting accuracy within the investigated sample compared to survey-based counterparts.

2.2. Potential pitfalls and keyword selection

The use of Internet search data is generally limited and subject to ambiguities: first, younger people are still more likely to go online than older people, and old people are more severely affected by unemployment. Moreover, the motivation and intention to enter a specific keyword in a search engine is rather unclear. It is inconclusive, whether the user is searching for his or her own purpose or for other reasons that are not directly related to their personal situation. For example, when entering the keyword “Volkswagen”, is the user searching for a new car of this brand because he or she plans to buy it in the near future? Or, is the user just searching for general information about the latest news concerning the Volkswagen emissions scandal? Accordingly, the use of Google data as a single predictor variable in a forecasting model should be treated with caution. Furthermore, using Google data as an additional predictor in a forecasting model might appear more robust.

Keyword selection is crucial when constructing an indicator from Google Trends. As mentioned in Choi and Varian (2012), the keyword selection should take into account what people would search for if they became unemployed or threatened with unemployment. Therefore, the searches should mainly cover two topics: first, what benefits are available to an unemployed person? And second, where can the unemployed apply for unemployment benefit and where can they find a new job? Another difficulty in choosing a keyword is separating the really unemployed people from those who are looking for a new job but who already have a job.

Consequently, in this study the keyword “unemployment” is used to obtain the Google search index. It should be noted that this keyword includes all searches in the Google search engine that contain this specific word. The chosen index also consists of queries such as “declare for unemployment”, “unemployment benefits” or “unemployment office”, to name a few possibilities. Therefore, the index is based on a wider range of search queries. This keyword fulfils the following condition:people at risk of unemployment can, for example, ask for general information from the employment office about possible unemployment benefits. If these people are actually unemployed, then they may use Google to find how or where to register as unemployed. When using a keyword such as “jobs”, these search intentions are also taken into account, but all persons who are already employed and only looking for another position are also considered. An index that includes these people may not be an appropriate indicator of the evolution of the unemployment rate.

3. Econometric framework

This section describes the econometric framework that will be used in this paper. First, the construction of unobserved latent factors from an unbalanced panel via PCA is described. These monthly factors are used as predictors in the forecast models and are intended to reflect the current state of the US economy. The MIDAS approach for combining low and high frequency variables in a regression model is then specified. Given that the Google and IC data are available weekly and the unemployment rate has a monthly frequency, the MIDAS approach can take into account the actual frequency of these higher frequency predictors to predict the low frequency target variable without loss of information, which would otherwise be the case if we aggregated the higher frequency variables. Then, the combination of the monthly estimated factors as predictors and the MIDAS approach with the weekly predictors is presented.

3.1. Estimation of factors and data irregularities

The unobserved latent factors that are used as predictors in the forecasting models are derived by principal components from a large set of macroeconomic variables. This dataset contains a number of candidate predictors for the unemployment rate. These factors serve the purpose of dimension reduction and are intended to represent the current state of the economy in the forecast models.

As specified in Stock and Watson (2002), let y_{t+1} be a time series that we wish to forecast and let X_t be an N-dimensional huge set of macroeconomic candidate predictor variables, as observed in $t = 1,\dots,T$. It is assumed that (X_t, y_{t+1}) admit a dynamic factor model representation with $\bar{r}$ common dynamic factors f_t,

$$y_{t+1} = \gamma(L)\, y_t + \beta(L)\, f_t + \varepsilon_{t+1}, \tag{3.1}$$

and

$$X_{it} = \lambda_i(L)\, f_t + e_{it}, \tag{3.2}$$

for $i = 1,\dots,N$, where $e_t = (e_{1t},\dots,e_{Nt})'$ is the $N \times 1$ idiosyncratic disturbance, and $\lambda_i(L)$ and $\beta(L)$ are lag polynomials. Stock and Watson (2002) make modifications to (3.1) and (3.2), namely the lag polynomials $\lambda_i(L)$, $\beta(L)$, and $\gamma(L)$ are modeled as having finite orders of q, so $\lambda_i(L) = \sum_{j=0}^{q} \lambda_{ij} L^j$ and $\beta(L) = \sum_{j=0}^{q} \beta_j L^j$. The finite lag assumption allows us to rewrite (3.1) and (3.2) in static form as

$$y_{t+1} = \gamma(L)\, y_t + \beta' F_t + \varepsilon_{t+1}, \tag{3.3}$$

and

$$X_t = \Lambda F_t + e_t, \tag{3.4}$$

where $F_t = \left(f_t', \ldots, f_{t-q}'\right)$ is $r \times 1$, where $r \leq (q+1)\,\bar{r}$, the ith row of Λ in (3.4) is $\left(\lambda_{i0}, \ldots, \lambda_{iq}\right)$, and $\beta = \left(\beta_0, \ldots, \beta_q\right)'$. This representation of the dynamic factor model enables the unobserved factors to be estimated by principal components.

However, estimating the static factors by principal components requires a balanced panel. When estimating macroeconomic variables in real-time, some data series have observations through the current period, whereas for others the most recent observations may only be available for a month or a quarter earlier. Therefore, the underlying dataset is unbalanced and the standard PCA is not applicable. The unbalanced feature of the data is called a "jagged edge" by Giannone et al. (2008). To cope with this jagged edge structure, Stock and Watson (2002) use the expectation-maximization (EM) algorithm.

As explained in McCracken and Ng (2016), the EM algorithm works as follows: observations that are missing are replaced by the unconditional mean based on the non-missing values, so that the panel is rebalanced. The unconditional mean is zero because the data are demeaned and standardized with mean zero before. A $T \times r$ matrix of factors $F = (f_1, \ldots, f_T)'$ and a $N \times r$ matrix of loadings $\lambda = (\lambda_1, \ldots, \lambda_N)$ are estimated from this balanced panel using the normalization that $\lambda'\lambda/N = I_r$. The missing value for series i at time t is updated from zero to $\widehat{\lambda}_i'\widehat{f}_t$. This value is multiplied by the standard deviation of the series and the mean is re-added. The resulting value is treated as an observation for series i at time t, and the mean and variance of the complete sample are re-calculated. The data are demeaned and standardized again, and the factors are re-estimated from this updated panel. The iteration stops when the factor estimates do not change.[3]

3.2. MIDAS estimation

3.2.1. MIDAS setup

Because the monthly unemployment rate and the weekly Google Trends data and the weekly number of initial claims are collected with different frequencies, the MIDAS regression proposed by Ghysels et al. (2006, 2007) is used. This represents a parsimonious class of time series models that allow us to capture the left-hand and right-hand variables of time series regressions with different frequencies.

As in Ghysels et al. (2007) and Andreou et al. (2011), we consider two variables to illus-

[3] For a detailed description of the use of the EM algorithm in combination with an unbalanced panel and data irregularities, see Stock and Watson (2002).

trate the MIDAS model: suppose that a variable y_{t+1} is observed once in a period, while data on a predictor variable $x_t^{(m)}$ are observed m times in between the same period. For example, y_{t+1} is the monthly unemployment rate, whereas $x_t^{(m)}$ is an adequate weekly predictor variable. Assuming that the number of weekly observations during a month is constant, then $m = 4$.

To generate direct multi-step-ahead forecasts, lagged values of the predictor variable $x_t^{(m)}$ are used to forecast y_{t+1}. The lagged values of $x_t^{(m)}$ are denoted by $x_{t-j/m}^{(m)}$, where the superscript on $x_{t-j/m}^{(m)}$ denotes the higher sampling frequency of $x_t^{(m)}$ to y_{t+1}.

When $h \geq 1$ denotes the h-step ahead forecast horizon, the general direct forecast MIDAS model, including lagged values of the higher-frequency predictor variable, can be written as

$$y_{\tau+h} = \beta_0 + B\left(L^{1/m};\theta\right)x_\tau^{(m)} + \varepsilon_{\tau+h}, \tag{3.5}$$

with $\tau = 1,\ldots,t-h$, where $B\left(L^{1/m};\theta\right) = \sum_{k=0}^{p_k-1} b(k;\theta)L^{k/m}$, and $L^{1/m}$ is a lag operator such that $L^{1/m}x_\tau^{(m)} = x_{\tau-1/m}^{(m)}$, and p_k is the maximum lag length of the predictors. The lag coefficients in $b(k;\theta)$ of the corresponding lag operator $L^{k/m}$ are parameterized as a function of a small-dimensional vector of parameters $\theta = (\theta_0,\theta_1,\ldots,\theta_p)$. This term acts as a weighting scheme that reduces the number of parameters to be estimated and leads to a more parsimonious model, instead of estimating a single coefficient for each high frequency lag. As stated in Pettenuzzo et al. (2016), the MIDAS model can also be written as

$$y_{\tau+h} = \beta_0 + \beta_1 B_1\left(L^{1/m};\theta_1\right)x_\tau^{(m)} + \varepsilon_{\tau+h}, \tag{3.6}$$

with $\tau = 1,\ldots,t-h$, where $\beta_1 B_1\left(L^{1/m};\theta_1\right) = B\left(L^{1/m};\theta\right)$. The weights are normalized to sum up to unity, so that the parameter β_1 captures the overall impact of the lagged values of $x_\tau^{(m)}$ on $y_{\tau+h}$.

3.2.2. MIDAS weighting function

The weighting scheme is crucial in MIDAS regression because it determines how the high-frequency predictor variable affects its regression coefficient in the MIDAS regression, and thus it determines its impact on the low-frequency target variable.

The parametrizations of the lagged coefficients of $b(k;\theta)$ proposed by Ghysels et al. (2007) can take various shapes for different values of θ. In general, the parameterized weights can decrease at different rates as the number of lags increases. By estimating θ, the given data selects the number of lags that are needed in the mixed-frequency-data relation between $y_{\tau+h}$

and x_τ.

As explained in Ghysels et al. (2007), several weighting schemes are available to reduce the number of parameters to be estimated. In this paper we use the two finite polynomials, the normalized exponential Almon lag polynomial and the normalized Beta probability density function. Andreou et al. (2010) show that these flexible weighting schemes, which are estimated by nonlinear least squares (NLS), are appropriate for forecasting purposes due to their parsimonious representation and flexible shapes.

Ghysels et al. (2007) state the normalized exponential Almon lag polynomial in the following general form:

$$b(k;\theta) = \frac{exp\left(\theta_1 k + \cdots + \theta_Q k^Q\right)}{\sum\limits_{k=1}^{K} exp\left(\theta_1 k + \cdots + \theta_Q k^Q\right)}. \tag{3.7}$$

The different shapes of the polynomials are only determined by the value of the parameters θ. Ghysels et al. (2005) use (3.7) with two parameters $\theta = [\theta_1;\theta_2]$. The resulting functional form is typically unimodal and can be slow-declining, fast-declining, hump shaped or flat. A declining shape implies that recent information receives a higher weight than earlier information. Accordingly, the rate of decline determines how many lags of the predictor variables are included in the regression model (3.5).

The normalized Beta probability density function as explained in Ghysels et al. (2007) also consists of two parameters $\theta = [\theta_1;\theta_2]$:

$$b(k;\theta_1,\theta_2) = \frac{f\left(\frac{k}{K},\theta_1;\theta_2\right)}{\sum\limits_{k=1}^{K} f\left(\frac{k}{K},\theta_1;\theta_2\right)}, \tag{3.8}$$

where:

$$f(x,a,b) = \frac{x^{a-1}(1-x)^{b-1}\Gamma(a+b)}{\Gamma(a)\Gamma(b)}, \tag{3.9}$$

with Γ as the gamma function

$$\Gamma(a) = \int_0^\infty e^{-x}x^{a-1}dx. \tag{3.10}$$

Similar to the normalized exponential Almon lag polynomial, the shape of the function is determined by the values of the parameter θ and the rate of decline decides how many lags are included in the regression model.[4]

[4]For a visual description of the functional forms of the MIDAS weights determined by the normalized exponential Almon lag polynomial and the normalized Beta probability density function, see Figures 5 and 6 in Appendix A.4.

3.3. FADL-MIDAS model

To forecast the low-frequency target variable, this paper uses a combination of the low-frequency estimated factors and the high-frequency MIDAS framework. Andreou et al. (2011) term this combination the factor augmented distributed lag MIDAS (FADL-MIDAS) model: a number of unobserved latent factors, which have the same frequency as the target variable, augment the MIDAS regression, where one or more high-frequency variables are used to predict the target variable.

To yield the general form of the FADL-MIDAS model, the model in (3.5) is extended with an autoregressive part of the target variable and with r estimated factors $F_t = (F_{1t}, \ldots, F_{rt})'$. Hence, the FADL-MIDAS model after Pettenuzzo et al. (2016) has the following representation:[5]

$$y_{\tau+h} = \alpha + \sum_{j=0}^{p_y - 1} \gamma_{j+1} y_{\tau-j} + \sum_{j=0}^{p_F - 1} \beta'_{j+1} F_{\tau-j} + B\left(L^{1/m};\theta\right) x_{\tau}^{(m)} + \varepsilon_{\tau+h}, \tag{3.11}$$

with $\tau = 1, \ldots, t-h$, and $B\left(L^{1/m};\theta\right) = \sum_{k=0}^{p_{k-1}} b(k;\theta) L^{k/m}$, where p_y and p_F denote the lag lengths of y_t and F_t, respectively. The lag length of the high-frequency predictors are represented by p_k. The functional form of the MIDAS weights $B\left(L^{1/m};\theta\right)$ depends either on the exponential Almon lag or on the normalized Beta function.

4. Data and forecasting design

4.1. Data

The data that is used in the empirical part comes from several sources and have several different frequencies: the unobserved latent factors are estimated from a large macroeconomic database with monthly time series, while Google Trends data and the number of initial claims are available weekly.

The predictor variable that is most noteworthy in this paper is the weekly GI. Google Trends is based on the Google web search engine and it provides a time series index that shows the volume of a particular search query or keyword, which is entered into the web search engine by Google users in a given geographic area within a given time. As stated in Choi and Varian (2012), the index is based on a share of the search queries: the number of web searches containing the keyword is normalized by dividing the total number of web searches performed through

[5] For further different representations of the FADL-MIDAS model, see Andreou et al. (2011) or Andreou et al. (2013).

Google for the same given time and region. The maximum query share in the time period is normalized to 100. The GI data are provided by Google to the public if the number of searches exceeds a certain unknown threshold. The data are available almost in real-time, starting with the first complete week in January 2004.[6] As explained in Section 2.2, the GI that we have used in the empirical exercise in this paper is obtained from the keyword "unemployment".

Given that the weekly GI has a particular seasonality, especially in November and December, when the total number of Google searches increases due to Christmas searches (D'Amuri and Marcucci, 2017), the time series is weekly seasonally adjusted with Seasonal-Trend decomposition procedure (STL). STL is capable of flexibly decomposing a high-frequent time series into trend, seasonal and remainder components based on Loess (Cleveland et al., 1990).[7]

The monthly unobserved latent factors are estimated from FRED-MD, which is the monthly database for Macroeconomic Research of the Federal Reserve Bank of St. Louis, which is described extensively in McCracken and Ng (2016). FRED-MD is a large macroeconomic database designed for the empirical analysis of "big data". The database is publicly available and updated on a monthly basis.[8] It consists of 134 monthly time series and is classified into eight categories: (1) output and income, (2) labor market, (3) housing, (4) consumption, orders and inventories, (5) money and credit, (6) interest and exchange rates, (7) prices and (8) stock market. A full list of the data and its transformation is given in Appendix A.2.

The target variable, which is the seasonally adjusted monthly US civilian unemployment rate, is released by the US BLS and is retrieved from FRED. The number of weekly initial claims, which is a more traditional predictor variable of the unemployment rate compared to Google data, is published by the US Employment and Training Administration of the US Department of Labor and is also retrieved from FRED.

The focus of this paper lies on short-term forecasts in real-time. Hence, all information that is given up to a certain date is used to conduct the forecasts.

Following D'Amuri and Marcucci (2017), the construction of the GI and the IC for month t is aligned with the time interval that is used to calculate the unemployment rate for month t, which is reported by the US BLS. Specifically, month t is defined by the week that includes the 12th of the corresponding month, the reference week, and the three preceding weeks. When

[6]Google Trends is available under the following link: `https://trends.google.com/trends/`. To conduct and download data about search queries via Google, a Google account is required.

[7]STL can be used in practice with most of the chosen parameters in an automated way. The seasonal smoothing parameter $n_{(s)}$ is chosen periodically and the trend smoothing parameter $n_{(t)}$ is calculated as follows: $n_{(t)} = \left[1.5 n_{(p)} / \left(1 - 1.5/n_{(s)}\right)\right]_{odd}$, where $n_{(p)}$ is the number of observations per seasonal cycle. For further details see Cleveland et al. (1990).

[8]The FRED-MD database is available for download under the following link: `https://research.stlouisfed.org/econ/mccracken/fred-databases/`.

there are more than four weeks between the reference week of month t and the following one in month $t+1$, the first week after the reference week is not used to calculate the unemployment rate.[9] Hence, this week is also excluded when calculating the GI and the IC in this paper.

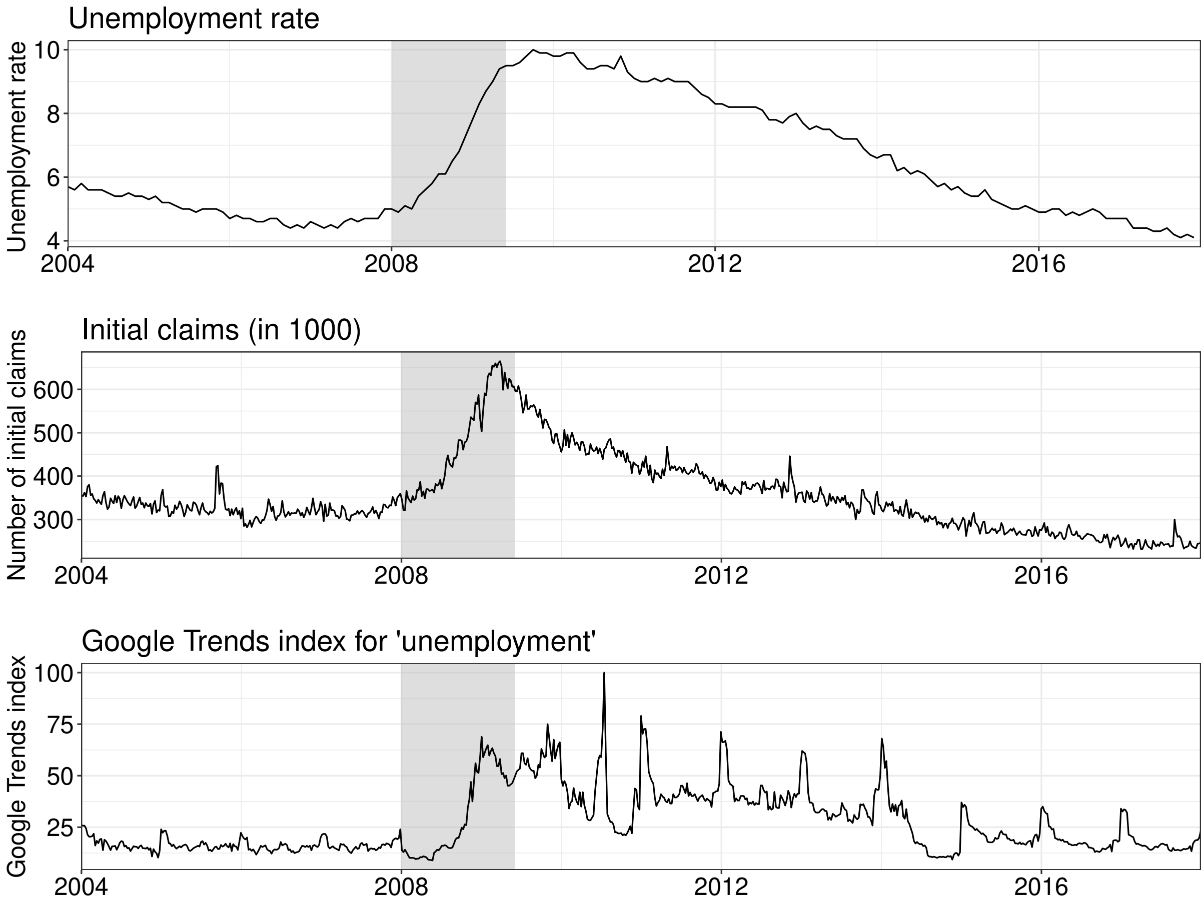
Figure 1 shows plots of the US unemployment rate, the IC and the GI.[10] It appears that the IC reached its peak during the 2008/09 recession[11], while the unemployment rate peaked outside the NBER recession band in October 2009. The GI experienced a very sudden sharp rise during the recession, but decreased again towards the end of the recession. After the recession, the index peaked in August 2010 and then fell over time with sharp fluctuations. It appears that the peaks of the GI after the recession correspond to the peaks of the IC and the unemployment rate. The IC tends to move towards the unemployment rate, while the GI also shows higher volatility and seasonality overall. This illustrates that both series are suitable early indicators for the unemployment rate and can thus be useful predictors in a forecasting framework.

[9] Figure 7 in Appendix A.5 shows a visual description of the exact timing of the unemployment rate calculation.

[10] The corresponding descriptive statistics are provided in Table 4 in Appendix A.1.

[11] The NBER dates the recession from 2008:M1 to 2009:M6. The NBER recession dates are obtained from FRED: `https://fred.stlouisfed.org/series/USREC`.

Figure 1: Time series plots.



Notes: The top panel shows the seasonally adjusted US unemployment rate in monthly frequency. The middle panel shows the seasonally adjusted number of initial claims in weekly frequency. The bottom panel shows the raw Google Trends index for the keyword "unemployment" in weekly frequency. The samples start in January 2004 and ending in December 2017. NBER recessions are highlighted by gray shading.

4.2. Forecasting setup and simulated real-time design

The empirical analysis in this paper focuses on the multi-step out-of-sample forecasting performance of three different DI variants proposed by Stock and Watson (2002) using the FADL-MIDAS forecasting model stated in (3.11) for forecasting the US unemployment rate. Following Stock and Watson (2002), the first DI variant is denoted by "DI-AR, lag": this includes $p_y = 5$ lags of the target variable and $p_F = 2$ lags of l monthly factors, where l is the number of estimated unobserved latent factors used in the regression. In addition, 12 weeks of corresponding weekly observations are included, hence $p_k = 12$. The second model, denoted "DI-AR", contains an autoregressive part with $p_y = 5$, l contemporaneous factors and four weeks of weekly information; therefore, $p_F = 0$ and $p_k = 4$. The third model, denoted "DI", includes only l contemporaneous monthly factors and four weeks of corresponding weekly information; therefore, $p_y = 0$, $p_F = 0$, and $p_k = 4$.

This paper focuses on short-term forecasting, so that the forecast horizon h is set to $h = 1,2,\dots,12$, meaning that one- to twelve-step-ahead direct forecasts of the monthly US unemployment rate are conducted.

The h-step-ahead out-of-sample forecasting performances of the three DI variants are then compared to the forecasting performance of a benchmark model.

The first benchmark model is a univariate AR(p) model based on (3.11), where p_F and p_k are set to zero. Hence, the general form of the benchmark model is:

$$y_{\tau+h} = \alpha + \sum_{j=0}^{p-1} \gamma_{j+1} y_{\tau-j} + \eta_{\tau+h}, \tag{4.1}$$

with $\tau = 1,\dots,t-h$, where the lag order p is selected recursively by the Akaike information criterion (AIC), with $0 \leq p \leq 4$.

The three DI variants and the AR(p)-benchmark are estimated by ordinary least squares (OLS) with a rolling window, whereby the window size is set to $S = 37$ similar to D'Amuri and Marcucci (2017). The whole estimation procedure is conducted recursively in real-time, as done in Stock and Watson (2002).

To construct these real-time forecasts, the data of the FRED-MD database are first screened for outliers. The data are then standardized with zero mean and unit variance, missing data are replaced by the EM, and then the factors are estimated by principal components. These factors are then checked for stationarity and are used in the three different DI forecasting variants, together with the seasonally adjusted GI and the IC as additional weekly regressors. Going one step further in time, all factors, the seasonally adjusted GI, the parameters, and so forth are then re-estimated and new forecasts are made.

Due to the constraint of the Google Trends data, the sample starts in 2004:M2. The first simulated out-of-sample forecast is made in 2008:M3. Therefore, the first one-step-ahead out-of-sample forecast is made for 2008:M4, while the first twelve-step-ahead out-of-sample forecast is made for 2009:M3. The final simulated out-of-sample forecast is made for 2017:M12.

Given that the GI has a short history and the forecasting models should be parsimonious concerning the number of predictors, the number of estimated factors l used in the forecasting models is set to four. In contrast, the PC_{p2} information criterion proposed by Bai and Ng (2002) finds eight factors in the FRED-MD database. However, using this number of factors would require many more observations in the forecasting models, so that the first forecast could only be made for a later date. Because one of the aims of this paper is to investigate if high-frequency data has better forecasting performance than the usual benchmark models, especially during the Great Recession in 2008/09, the first forecast should be made for 2008 and not later. Hence, the forecasting models have to be parsimonious concerning the number of predictors and the required observations. Consequently, the number of factors is reduced to four. In addition, using too many factors might lead to overfitting and this would result in a poor forecasting performance. Figure 4 in Appendix A.3 shows a Scree plot of the estimated factors and Table 5 in Appendix A.3 illustrates that the first four factors explain 34 % of the variation of the FRED-MD data, while eight factors explain 47,5 % of the variation.

5. Empirical results

5.1. Methodology

The out-of-sample forecasts of the three different DI variants are first evaluated relative to the AR(p)-benchmark by comparing the RMSFE of each DI model with that of the benchmark model. The benchmark model has a relative RMSFE of one, whereas a value below one indicates that the competitor model has a lower RMSFE than the benchmark model and thus outperforms it. The results of the Diebold and Mariano (1995) (DM) test are also presented. The DM test generally tests the null hypothesis of no difference in the accuracy of two competing forecasts. The relative predictive performances of the DI variants used in this paper are additionally compared with the forecasts made in D'Amuri and Marcucci (2017) from a similar forecasting framework, where the authors use monthly averages of a Google index formed from the keyword "jobs"[12].

Following D'Amuri and Marcucci (2017), the forecasts are also visually compared by the cumulative sum of squared forecast error differences (CSSED) introduced by Welch and Goyal (2008). A visual representation of the CSSED allows a quick and simple overview to decide, whether a benchmark model is outperformed by a competitor model. As stated in D'Amuri and Marcucci (2017), the CSSED is computed as $CSSED_{m,\tau} = \sum_{\tau=R}^{T}\left(\hat{e}^2_{bm,\tau} - \hat{e}^2_{m,\tau}\right)$, where $\hat{e}^2_{bm,\tau}$ denotes the squared forecast error of a benchmark model. The squared forecast error of the competitor model is denoted by $\hat{e}^2_{m,\tau}$, and R and T indicate the beginning and end of the

[12]To be more precise, D'Amuri and Marcucci (2017) subtract the keyword "Steve Jobs" from the keyword "jobs" to improve the precision of their index.

forecast evaluation sample. A CSSED value above zero generally indicates a better forecasting performance compared to the benchmark model at this point. Positive changes in the slope of the CSSED lines indicate that the relative predictive performance of competing models increases compared to the benchmark model, while negative changes represent a decrease in relative performance.

Moreover, for each of the three different DI variants, in total 12 different forecasts are made. On the one hand, either the Almon lag or the Beta function are used as specification weights for the weekly GI and IC. On the other hand, the predictors are considered differently in the forecasting models based on the FADL-MIDAS model from (3.11) to get further insight into which combination of predictor variables might give the best results. To be precise, in the first model, only the GI data are considered, whereas the factors and the IC are excluded from the model. In the second model, the forecasts are based on the IC, whereas the GI and the factors are excluded. The third model uses the GI and IC together, with the fourth and fifth models using the individual indicators together with the factors. The sixth model is the largest and it considers all three predictors.

5.2. Comparison with AR(p)-benchmark model

Table 1 shows the monthly out-of-sample forecast results for the US unemployment rate compared to forecasts from the AR(p)-benchmark. For each forecast horizon h, the RMSFE of the AR(p)-benchmark is outperformed but there is no clear pattern to indicate which forecast model is best overall because the model that combines the lowest RMSFE changes from horizon to horizon. It should be noted that the DM test does not find statistical significance for the shortest forecast horizon $h = 1$, but for all other remaining forecast horizons, so that for these horizons the forecast accuracy of the models presented in this paper is statistically better than the AR(p)-benchmark.

In view of the one-step-ahead forecasts ($h = 1$), the 'DI-AR, lag' model with the Almon lag as weighting scheme has the lowest overall relative RMSFE compared to the AR(p)-benchmark model. It is worth mentioning here that for many models the Beta function and also partly the Almon lag select the same lags and coefficients, so that the final results are the same. For the shortest forecast horizons, GI and IC provide the best forecast results as individual predictor variables or as a combination, whereas for longer forecast horizons, the combination of factors and IC provides the best results. For forecasting horizons of three months and more, the forecasting models in which the factors play a role as indicators also have, for the most part, a better forecasting capability than the benchmark, which means that the factors are not suitable for the shortest but are suitable for a somewhat longer period of time as indicators.

Figure 2 shows the development of CSSED values for each forecast horizon h. In each plot, only the models with the lowest RMSFE for each panel per forecast horizon from Table 1 are plotted, so that the following discussion is limited to these models only. In all plots, the AR(p)-benchmark is defined as a horizontal line with intercept zero. Because the first simulated out-of-sample forecast is performed in 2008:M3, the CSSED chart begins immediately after the start of the 2008/09 recession.

Taking into account the CSSED chart for the forecast horizon $h = 1$, only the best Panel A, B and C models have a value above zero for the entire sample, outperforming the benchmark at all times, with Panel A having the highest CSSED value at the end of the sample, which explains the final result shown in Table 1. The best Panel D and F models show a mainly positive increase during the recession, but decrease in value at the end of the recession and continue to decline in total over time.

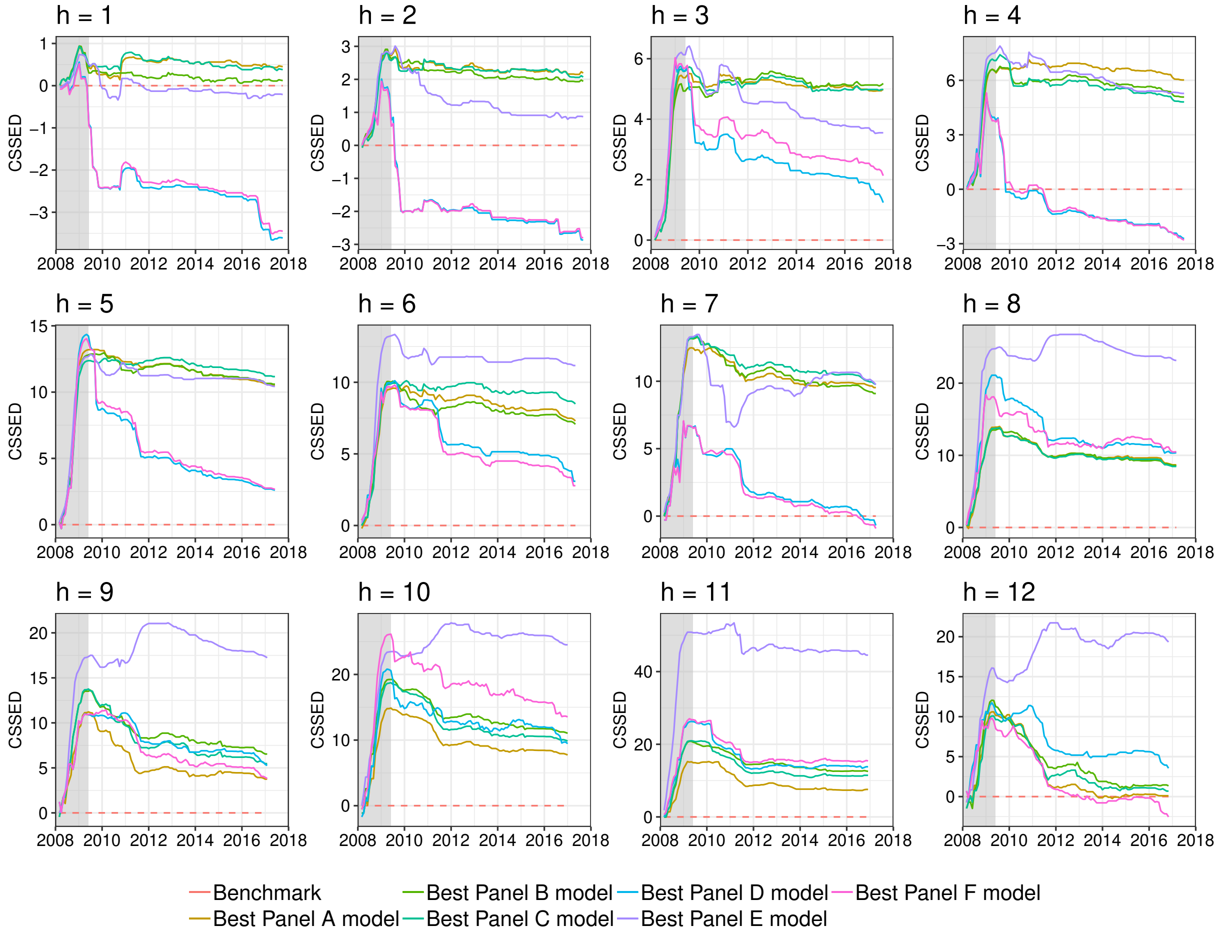
For the remaining plots, the behaviors of the CSSED lines are generally quite similar. During the recession, the CSSED lines show a strong rise, which shows that the competing models clearly outperform the AR(p)-benchmark at this time. When the recession comes to an end, the CSSED lines tend to stabilise and then fall more or less over time depending on the model, with by far the majority of models remaining at a value above zero. These visualisations show that all of the forecasting models presented in this article achieve their better overall forecasting performance for the entire sample compared to the AR(p)-benchmark during the 2008/09 recession. For the longer forecast horizons, $h = 6$ to $h = 12$, the combination of factors and IC reaches the highest CSSED values, which indicates that this combination is by far the most promising for the longest forecast horizons presented in this paper. With one exception at horizon $h = 7$, the CSSED values of the best Panel E models stabilise and improve unlike the other models presented, even after the end of the recession.

Table 1: Out-of-sample forecast results for the monthly US unemployment rate with AR(p)-benchmark.

model	*weighting scheme*	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
Panel A: Google Trends index													
DI-AR, lag	Almon	**0.946**	**0.878****	*0.852****	**0.891*****	0.878**	*0.937***	*0.938***	0.960*	*0.985*	0.977*	*0.977**	*1.000*
	Beta	0.976	0.890**	0.876**	0.926**	*0.869***	0.946**	0.939**	*0.956***	0.989	*0.973**	0.978	1.015
DI-AR	Almon	0.976	0.890**	0.876**	0.926**	*0.869****	0.946***	0.939***	*0.956***	0.989	*0.973***	0.978*	1.015
	Beta	0.976	0.890**	0.876**	0.926**	*0.869***	0.946**	0.939**	*0.956***	0.989	*0.973**	0.978	1.015
DI	Almon	1.034	1.012	1.015	1.007	1.006	1.005	1.000	1.007	1.005	1.013	1.025	1.033
	Beta	1.034	1.012	1.015	1.007	1.006	1.005	1.000	1.007	1.005	1.013	1.025	1.033
Panel B: Initial claims													
DI-AR, lag	Almon	0.996	0.908**	0.848**	0.925**	*0.868***	0.940**	0.942**	*0.957***	*0.973*	0.965*	0.977*	1.001
	Beta	*0.986*	*0.893***	0.877**	0.915**	0.879**	*0.939***	0.942**	0.961**	0.984	0.964*	0.974	1.003
DI-AR	Almon	1.023	0.899*	**0.846*****	*0.909****	0.874***	0.942***	*0.941***	0.958**	0.980*	*0.961***	*0.961***	*0.996*
	Beta	*0.986*	*0.893***	0.877**	0.915**	0.879**	*0.939***	0.942**	0.961**	0.984	0.964*	0.974	1.003
DI	Almon	0.992	1.045	1.023	1.024	1.035	1.005	1.020	1.011	1.012	1.018	1.005	1.016
	Beta	1.037	1.070	1.053	1.048	1.033	1.027	1.024	1.019	1.019	1.019	1.016	1.016
Panel C: Google Trends index and initial claims													
DI-AR, lag	Almon	*0.954*	*0.885**	0.866**	0.931**	**0.860*****	*0.927***	0.943**	0.967*	*0.977*	0.967*	0.979*	*0.998*
	Beta	1.001	0.892*	0.859**	0.918*	0.882**	0.939**	**0.937****	*0.957***	0.985	0.968*	0.978	1.012
DI-AR	Almon	1.005	0.908*	*0.853****	*0.914***	0.873***	0.934***	0.940**	0.959**	0.992	*0.965**	*0.965***	1.011
	Beta	1.001	0.892*	0.859**	0.918*	0.882**	0.939**	**0.937****	*0.957***	0.985	0.968*	0.978	1.012
DI	Almon	1.031	1.085	1.019	1.018	1.033	1.002	1.021	1.008	1.009	1.011	1.006	1.018
	Beta	0.988	1.012	0.991	1.011	1.005	1.003	1.004	1.008	1.013	1.021	1.027	1.040
Panel D: Factors and Google Trends index													
DI-AR, lag	Almon	1.929	1.431	1.078	1.227	1.115	1.021	1.098	0.967	1.078	0.970	0.995	1.057
	Beta	1.970	1.412	1.071	1.187	1.129	1.048	1.134	0.982	1.076	*0.967*	1.022	1.066
DI-AR	Almon	1.531	1.215	0.993	1.110	*0.969*	1.022	1.026	*0.947**	1.011	0.970*	*0.957**	1.026
	Beta	1.531	1.215	0.993	1.110	*0.969*	1.022	1.026	*0.947**	1.011	0.970*	*0.957**	1.026
DI	Almon	1.377	1.148	*0.965*	*1.046*	1.006	*0.974*	*1.004*	0.953*	*0.978*	0.979	0.973	*0.990*
	Beta	*1.360*	*1.140*	0.969	1.050	1.012	0.978	1.008	0.958	0.981	0.984	0.978	0.999
Panel E: Factors and initial claims													
DI-AR, lag	Almon	2.371	1.366	1.154	1.198	1.192	1.094	1.157	0.988	1.043	0.953*	1.043	1.097
	Beta	1.536	1.391	1.074	1.107	1.039	1.072	1.109	1.006	0.997	1.040	**0.854*****	0.972
DI-AR	Almon	1.440	1.204	0.966	1.109	0.968	1.031	1.028	0.942*	1.014	0.964*	0.952**	1.017
	Beta	1.117	1.018	0.914*	0.991	0.888**	0.986	**0.937****	0.951**	0.930***	**0.911*****	0.888***	**0.946*****
DI	Almon	1.457	1.085	0.957	1.035	0.983	0.982	0.998	0.957	0.983	0.974	0.966*	0.989
	Beta	*1.024*	*0.953*	*0.898**	*0.905***	*0.870****	**0.902*****	0.951***	**0.878*****	**0.927*****	0.920***	0.863***	0.981*
Panel F: Factors, Google Trends index and initial claims													
DI-AR, lag	Almon	2.188	1.484	1.136	1.134	1.173	1.106	1.121	0.972	1.020	0.985	0.993	1.049
	Beta	2.209	1.491	1.348	1.137	1.130	1.079	1.138	*0.946**	1.024	*0.952**	0.992	1.022
DI-AR	Almon	1.637	1.299	0.948	1.117	0.969	1.036	1.031	*0.946**	1.013	0.971	*0.952**	1.032
	Beta	1.535	1.250	0.985	1.133	*0.968*	1.017	1.029	0.957*	1.011	0.966*	0.959*	1.033
DI	Almon	1.375	1.160	*0.940*	*1.047*	1.013	0.993	*1.006*	0.954	*0.984*	0.979	0.975	*1.007*
	Beta	*1.346*	*1.137*	0.953	1.067	1.007	*0.977*	1.008	0.957	0.988	0.982	0.980	1.010

Notes: The table reports the relative RMSFE of the competitor model to the AR(p)-benchmark model. A value below one indicates that the competitor model beats the benchmark model and vice versa. The numbers in bold correspond to the lowest RMSFE at each forecast horizon h, whereas the values in italic are the lowest RMSFE for each panel at each forecast horizon h. The stars denote statistical significance at $10\,\%(^{*})$, $5\,\%(^{**})$ and $1\,\%(^{***})$ level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the AR(p) benchmark and the competitor model when the benchmark is beaten.

Figure 2: CSSED comparison plots with AR(p)-benchmark model.



Notes: NBER recessions are highlighted by gray shading.

5.3. Comparison with D'Amuri and Marcucci (2017)-benchmark model

Table 2 shows the monthly out-of-sample forecast results for the US unemployment rate compared to the forecasts made using the approach presented in D'Amuri and Marcucci (2017).

Similar to the comparison with the AR(p)-benchmark, there is no clear pattern as to which forecast model is best overall because the model that combines the lowest RMSFE changes from horizon to horizon.

For the first two horizons, the benchmark can be slightly outperformed by the GI as a single predictor variable but the DM test does not find any significant difference compared to the benchmark. For $h = 3$, the use of the IC as single predictor variable gives the overall best result, whereas the GI and the combination of the IC and GI also have a lower RMSFE than the benchmark. For the forecasts for the next four and five months, the benchmark cannot be exceeded by any of the models presented in this paper. However, for the forecast horizons $h = 6$ and $h = 7$ the benchmark is again exceeded and especially for the forecast in six months the DM test finds a significant prediction accuracy compared to the benchmark at the level of 1 % for many of the models presented, with the lowest RMSFE for the DI variant with the combination of factors and the IC as predictor variables. For the remaining horizons, the benchmark cannot be exceeded.

The corresponding CSSED plots in Figure 3 illustrate these findings, with the horizontal representing the D'Amuri and Marcucci (2017)-benchmark.

For the one-month-ahead forecasts, the CSSED values of the best models rise during the recession and continue to rise until 2011, and then stabilise thereafter. Therefore, a forecast gain can also be determined after the end of the 2008/09 recession compared to the benchmark. For $h = 2$ to $h = 5$, most models lose their different forecast gains against the benchmark over time after the recession and, in particular, for $h = 4$ and $h = 5$, all CSSED values finally assume negative values over time. All six-months-ahead forecasts clearly outperform the benchmark. The performance increases during the recession and most models continue to improve over time. At $h = 7$, the forecast performance of the various models varies greatly over time, with the best Panel D and F models, for example, showing a forecast gain over the benchmark in the first months of the recession but then declining very sharply in value over time with fluctuations. For the horizons $h = 9$ to $h = 12$, none of the competitor models outperforms the benchmark at any point in time, and the performance declines for the most part over time.

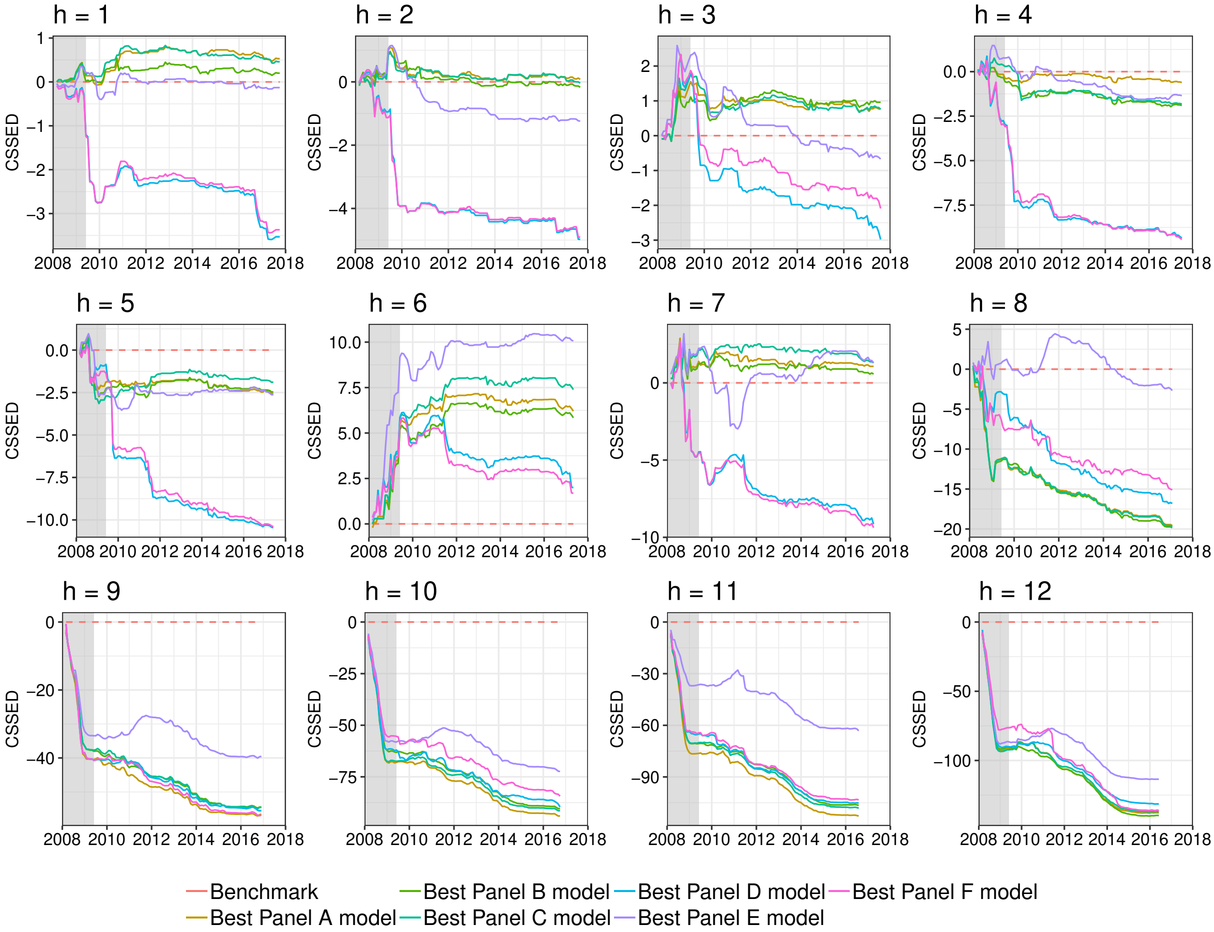
The results show that the forecasting framework presented in this paper with the MIDAS approach and the selection of predictors has some advantages over the D'Amuri and Marcucci (2017) approach in the short-term, with the exception of forecasting horizons of four and five months, but produces worse results in the longer forecasting period, namely 8- to 12-months-ahead. It can be generally stated that the quality of the forecast is subject to fluctuations over the different forecast horizons in comparison to this benchmark.

Table 2: Out-of-sample forecast results for the monthly US unemployment rate with D'Amuri and Marcucci (2017)-benchmark.

model	*weighting scheme*	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
Panel A: Google Trends index													
DI-AR, lag	Almon	**0.937**	**0.995**	*0.972*	**1.013**	1.053	*0.946***	*0.992*	1.103	*1.243*	1.392	*1.418*	*1.477*
	Beta	0.967	1.009	0.999	1.053	*1.042*	0.955**	0.993	*1.099*	1.251	*1.388*	1.423	1.503
DI-AR	Almon	0.967	1.009	0.999	1.053	*1.042*	0.955**	0.993	*1.099*	1.251	*1.388*	1.423	1.503
	Beta	0.967	1.009	0.999	1.053	*1.042*	0.955**	0.993	*1.099*	1.251	*1.388*	1.423	1.503
DI	Almon	1.024	1.146	1.157	1.146	1.206	1.014	1.058	1.146	1.265	1.432	1.469	1.515
	Beta	1.024	1.146	1.157	1.146	1.206	1.014	1.058	1.146	1.265	1.432	1.469	1.515
Panel B: Initial claims													
DI-AR, lag	Almon	0.987	1.028	0.967	1.052	*1.040*	0.949**	*0.996*	*1.100*	*1.235*	1.379	1.419	1.488
	Beta	*0.977*	*1.011*	1.001	*1.041*	1.054	*0.948***	*0.996*	1.102	1.247	1.380	1.415	1.487
DI-AR	Almon	1.014	1.019	**0.965**	1.034	1.048	0.951**	*0.996*	1.101	1.242	*1.376*	*1.399*	*1.483*
	Beta	*0.977*	*1.011*	1.001	*1.041*	1.054	*0.948***	0.996	1.102	1.247	1.380	1.415	1.487
DI	Almon	0.983	1.183	1.167	1.166	1.241	1.015	1.079	1.151	1.274	1.434	1.441	1.495
	Beta	1.027	1.212	1.201	1.193	1.239	1.036	1.083	1.159	1.281	1.434	1.455	1.495
Panel C: Google Trends index and initial claims													
DI-AR, lag	Almon	*0.945*	*1.003*	0.987	1.059	**1.031**	*0.935***	0.998	1.110	*1.235*	*1.379*	1.418	*1.474*
	Beta	0.992	1.010	0.979	1.045	1.057	0.947**	**0.991**	*1.099*	1.246	1.386	1.422	1.496
DI-AR	Almon	0.995	1.028	*0.973*	*1.040*	1.046	0.942**	0.994	1.100	1.254	1.380	*1.404*	1.499
	Beta	0.992	1.010	0.979	1.045	1.057	0.947**	**0.991**	*1.099*	1.246	1.386	1.422	1.496
DI	Almon	1.022	1.229	1.162	1.158	1.239	1.011	1.080	1.148	1.271	1.425	1.442	1.496
	Beta	0.979	1.147	1.130	1.150	1.205	1.012	1.062	1.149	1.274	1.442	1.472	1.522
Panel D: Factors and Google Trends index													
DI-AR, lag	Almon	1.911	1.621	1.229	1.396	1.337	1.030	1.161	1.092	1.318	1.376	1.420	1.495
	Beta	1.952	1.599	1.221	1.351	1.354	1.058	1.199	1.104	1.319	*1.373*	1.460	1.519
DI-AR	Almon	1.517	1.376	1.133	1.264	*1.162*	1.032	1.085	*1.085*	1.271	1.378	*1.395*	1.503
	Beta	1.517	1.376	1.133	1.264	*1.162*	1.032	1.085	*1.085*	1.271	1.378	*1.395*	1.503
DI	Almon	1.364	1.301	*1.101*	*1.190*	1.206	*0.983*	*1.062*	1.092	*1.239*	1.398	1.410	*1.458*
	Beta	*1.347*	*1.292*	1.106	1.195	1.213	0.987	1.066	1.096	1.240	1.403	1.420	1.474
Panel E: Factors and initial claims													
DI-AR, lag	Almon	2.349	1.547	1.316	1.363	1.429	1.104	1.224	1.121	1.290	1.356	1.483	1.578
	Beta	1.522	1.576	1.224	1.260	1.245	1.082	1.173	1.144	1.250	1.441	**1.252**	1.444
DI-AR	Almon	1.426	1.364	1.102	1.262	1.160	1.041	1.088	1.082	1.278	1.374	1.381	1.494
	Beta	1.106	1.153	1.043	1.128	1.065	0.995	**0.991**	1.088	1.183	**1.309**	1.291	**1.405**
DI	Almon	1.443	1.229	1.091	1.177	1.178	0.991	1.055	1.098	1.248	1.393	1.402	1.465
	Beta	*1.015*	*1.080*	*1.024*	*1.030*	*1.043*	**0.911*****	1.006	**1.014**	**1.175**	1.312	1.265	1.437
Panel F: Factors, Google Trends index and initial claims													
DI-AR, lag	Almon	2.168	1.681	1.296	1.290	1.406	1.116	1.185	1.101	1.260	1.391	1.418	1.490
	Beta	2.189	1.689	1.537	1.294	1.354	1.089	1.203	*1.077*	1.275	*1.354*	1.416	*1.472*
DI-AR	Almon	1.622	1.472	1.081	1.271	1.162	1.046	1.091	1.084	1.273	1.378	*1.389*	1.510
	Beta	1.521	1.416	1.123	1.290	*1.161*	1.026	1.089	1.096	1.271	1.375	1.395	1.508
DI	Almon	1.362	1.314	*1.072*	*1.191*	1.215	1.002	*1.064*	1.094	*1.244*	1.399	1.415	1.482
	Beta	*1.334*	*1.288*	1.087	1.214	1.207	*0.986*	1.066	1.096	1.246	1.402	1.421	1.484

Notes: The table reports the relative RMSFE of the competitor model to the benchmark forecasts after D'Amuri and Marcucci (2017). A value below one indicates that the competitor model beats the benchmark model and vice versa. The numbers in bold correspond to the lowest RMSFE at each forecast horizon h, whereas the values in italic are the lowest RMSFE for each panel at each forecast horizon h. The stars denote statistical significance at $10\%(^{*})$, $5\%(^{**})$ and $1\%(^{***})$ level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark model and the competitor model when the benchmark is beaten.

Figure 3: CSSED comparison plots with D'Amuri and Marcucci (2017)-benchmark.



Notes: NBER recessions are highlighted by gray shading.

To gain a deeper insight into the quality of the individual forecasts, the relative RMSFE of the forecast models for each individual month are determined and compared with the D'Amuri and Marcucci (2017)-benchmark. In the following, only the most promising models are discussed, and this only for the shortest forecast horizon $h = 1$.[13] Table 3 shows the results for the GI and IC as individual predictor variables and both predictors combined.

Once again, it turns out that no single model consistently produces the best results. It is shown that except for June, the relative RMSFE obtained for most models is below one, so the benchmark for these months is outperformed. However, the DM test does not always find a significant difference to the benchmark. This could happen due to the fact that the sample size decreases very sharply when the sample is broken down into individual months, so that

[13]The remaining results can be found in Appendix A.6. A discussion of all presented results would go beyond the scope of this paper.

each individual DM test is based on only very few values and the meaningfulness is therefore limited. In addition, these results and the further results listed in Appendix A.6 show that a kind of seasonality prevails in the forecasts compared to the benchmark because January, June and July are the months in which the benchmark is most difficult to outperform.

In conclusion, it can be said that the use of weekly data from the GI chosen in this paper and the number of initial claims in the short-term forecasting exercise mostly outperform the benchmark, although statistical significance cannot be determined for every point in time.

Table 3: One-step-ahead out-of-sample forecast results for each individual month for the US unemployment rate with D'Amuri and Marcucci (2017)-benchmark.

model	*weighting scheme*	*Jan.*	*Feb.*	*Mar.*	*Apr.*	*May*	*Jun.*	*Jul.*	*Aug.*	*Sept.*	*Oct.*	*Nov.*	*Dec.*
Panel A: Google Trends index													
DI-AR, lag	Almon	**0.891**	*0.612*	1.047	0.743**	0.874	1.225	1.472	*0.725***	0.808	*0.894*	1.028	**0.894**
	Beta	1.083	0.685	1.047	0.766**	*0.863*	1.279	1.429	*0.725***	*0.752*	0.922	1.028	0.992
DI-AR	Almon	1.083	0.685	1.047	0.766**	*0.863*	1.279	1.429	*0.725***	*0.752*	0.922	1.028	0.992
	Beta	1.083	0.685	1.047	0.766**	*0.863*	1.279	1.429	*0.725***	*0.752*	0.922	1.028	0.992
DI	Almon	1.402	1.046	*0.976*	*0.657*	0.934	*1.128*	*1.225*	1.026	0.933	1.118	*0.978*	1.103
	Beta	1.402	1.046	*0.976*	*0.657*	0.934	*1.128*	*1.225*	1.026	0.933	1.118	*0.978*	1.103
Panel B: Initial claims													
DI-AR, lag	Almon	*1.017*	0.586*	**0.724**	0.743**	**0.786**	1.414	1.041	**0.607****	0.956	0.922	0.941	1.291
	Beta	1.130	**0.433***	0.951	0.754**	0.863	1.552	1.291	0.649**	*0.722**	**0.671****	0.870	1.155
DI-AR	Almon	1.189	0.661	1.000	0.777*	0.863	1.297	1.384	0.688**	0.909	0.866	0.854	1.176
	Beta	1.130	**0.433***	0.951	0.754**	0.863	1.552	1.291	0.649**	*0.722**	**0.671****	0.870	1.155
DI	Almon	1.326	1.132	0.900	0.682*	0.991	1.168	**0.816**	0.946	1.043	1.245	*0.707*	*0.975*
	Beta	1.259	1.016	1.024	**0.587***	1.243	**1.108**	0.890	0.889	1.103	1.304	0.746	1.111
Panel C: Google Trends index and initial claims													
DI-AR, lag	Almon	*0.947*	*0.586*	0.976	0.754**	*0.820*	1.225	1.541	**0.607****	0.752	1.072	1.028	*0.940*
	Beta	1.287	0.661	0.926	0.777**	0.934	1.523	1.384	0.688**	**0.626****	*0.975*	1.021	0.966
DI-AR	Almon	1.339	0.750	1.024	0.754**	0.842	1.348	1.486	0.725*	0.885	1.025	0.986	0.957
	Beta	1.287	0.661	0.926	0.777**	0.934	1.523	1.384	0.688**	**0.626****	*0.975*	1.021	0.966
DI	Almon	1.565	1.016	0.976	0.809	1.000	1.279	*0.957*	0.973	1.022	1.342	**0.697**	1.041
	Beta	1.339	0.968	*0.787*	*0.743*	0.915	*1.128*	1.155	1.051	0.909	1.140	0.949	*0.940*

Notes: The table reports the relative RMSFE for the one-step-ahead forecasts for each individual month of the competitor model compared to the benchmark forecasts after D'Amuri and Marcucci (2017). The numbers in bold correspond to the lowest RMSFE for each individual month, whereas the values in italic are the lowest RMSFE for each panel for each individual month. The stars denote statistical significance at 10%(*), 5%(**) and 1%(***) level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

6. Conclusion

Given the recent interest in using Internet search data to forecast economic variables, the aim of this paper is to assess whether the use of Google search data is useful to forecast the US unemployment rate when compared with other more traditional predictor variables.

One of the difficulties of obtaining a reliable Google Trends index is the choice of the keyword. In this paper the keyword "unemployment" is chosen to obtain a predictor variable for the unemployment rate. The keyword should reflect what people look for when they become unemployed or threatened by unemployment. In addition, the keyword must separate the unemployed from those who are looking for a new job but who already have a job.

Since Google data also have some drawbacks, the number of initial claims are also considered as an additional predictor variable. To incorporate the current state of the US economy, unobserved latent factors are derived by principal components from a macroeconomic database. Because the Google index and the number of initial claims have a weekly frequency, while the US unemployment rate and also the derived factors have a monthly frequency, the MIDAS regression model is used to maintain the original data frequency.

Based on Stock and Watson (2002), three different diffusion index variants of the MIDAS approach are used. Using a real-time forecast design across the entire sample, the selected Google Trends index provides the best forecast results as a single predictor variable within the 'DI-AR, lag' variant and Almon lag polynomial as weighting scheme for the shortest forecast horizons $h = 1$ and $h = 2$. Unfortunately, statistically, it does not exceed the D'Amuri and Marcucci (2017) benchmark after the DM test, which has already proven the predictive power of an alternative monthly Google Trends index. Nevertheless, this illustrates that the use of the GI and also the IC applied here with weekly information in a MIDAS framework leads to a forecast gain for the shortest horizons, although a kind of seasonality against the benchmark prevails in the forecasts—as the analysis of the individual months has shown.

With increasing forecasting horizons, the forecasting performance of the best diffusion index variants decreases over time, indicating that the forecasting methods proposed in this paper are only useful for short-term forecasts. This can also be seen in the development of the CSSED lines. At the beginning of the sample during the 2008/09 recession, the forecast performance rises relative to the benchmarks and it then falls at different rates over time. In addition, the results provide a mixed picture of which predictor variables are most useful because the best model and combination of predictor variables change from horizon to horizon.

In summary, there is no clear indication as to which model and which predictor variables are best suited to forecast the US unemployment rate in this MIDAS framework and in this sample. Therefore, a combination of the best models should be used to get an idea of how the unemployment rate will behave in the near future. It will be interesting to see how the proposed forecasting models will behave as more data become available over time, especially in times of economic uncertainty. If the results are confirmed, then the use of Internet data alongside more traditional predictor variables could be routinely included in the near future.

References

Andreou, E., Ghysels, E., and Kourtellos, A. (2010). Regression models with mixed sampling frequencies. *Journal of Econometrics*, 158(2):246–261.

Andreou, E., Ghysels, E., and Kourtellos, A. (2011). Forecasting with mixed-frequency data. *Oxford handbook of economic forecasting*, pages 225–245.

Andreou, E., Ghysels, E., and Kourtellos, A. (2013). Should macroeconomic forecasters use daily financial data and how? *Journal of Business & Economic Statistics*, 31(2):240–251.

Askitas, N. and Zimmermann, K. F. (2009). Google econometrics and unemployment forecasting. *Applied Economics Quarterly*, 55(2):107–120.

Bai, J. and Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70(1):191–221.

Bangwayo-Skeete, P. F. and Skeete, R. W. (2015). Can Google data improve the forecasting performance of tourist arrivals? Mixed-data sampling approach. *Tourism Management*, 46:454–464.

Chen, T., So, E. P. K., Wu, L., and Yan, I. K. M. (2015). The 2007–2008 US recession: what did the real-time Google Trends data tell the United States? *Contemporary Economic Policy*, 33(2):395–403.

Choi, H. and Varian, H. (2012). Predicting the present with Google Trends. *Economic Record*, 88(1):2–9.

Cleveland, R. B., Cleveland, W. S., McRae, J. E., and Terpenning, I. (1990). STL: A Seasonal-Trend Decomposition Procedure based on Loess. *Journal of Official Statistics*, 6(1):3–73.

D'Amuri, F. and Marcucci, J. (2017). The predictive power of Google searches in forecasting US unemployment. *International Journal of Forecasting*, 33(4):801–816.

Dergiades, T., Milas, C., and Panagiotidis, T. (2014). Tweets, Google Trends, and sovereign spreads in the GIIPS. *Oxford Economic Papers*, 67(2):406–432.

Diebold, F. X. and Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3).

Fondeur, Y. and Karamé, F. (2013). Can Google data help predict French youth unemployment? *Economic Modelling*, 30:117–125.

Ghysels, E., Santa-Clara, P., and Valkanov, R. (2005). There is a risk-return trade-off after all. *Journal of Financial Economics*, 76(3):509–548.

Ghysels, E., Santa-Clara, P., and Valkanov, R. (2006). Predicting volatility: getting the most out of return data sampled at different frequencies. *Journal of Econometrics*, 131(1-2):59–95.

Ghysels, E., Sinko, A., and Valkanov, R. (2007). MIDAS regressions: Further results and new directions. *Econometric Reviews*, 26(1):53–90.

Giannone, D., Reichlin, L., and Small, D. (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, 55(4):665–676.

Guzmán, G. (2011). Internet search behavior as an economic forecasting tool: The case of inflation expectations. *Journal of economic and social measurement*, 36(3):119–167.

Liu, T., Xu, X., and Fan, F. (2018). Forecasting Chinese GDP using online data. *Emerging Markets Finance and Trade*, 54(4):733–746.

McCracken, M. W. and Ng, S. (2016). FRED-MD: A monthly database for macroeconomic research. *Journal of Business & Economic Statistics*, 34(4):574–589.

McLaren, N. and Shanbhogue, R. (2011). Using internet search data as economic indicators. *Bank of England Quarterly Bulletin*, 51(2).

Montgomery, A. L., Zarnowitz, V., Tsay, R. S., and Tiao, G. C. (1998). Forecasting the US unemployment rate. *Journal of the American Statistical Association*, 93(442):478–493.

Pettenuzzo, D., Timmermann, A., and Valkanov, R. (2016). A MIDAS approach to modeling first and second moment dynamics. *Journal of Econometrics*, 193(2):315–334.

Smith, P. (2016). Google's MIDAS touch: Predicting UK unemployment with internet search data. *Journal of Forecasting*, 35(3):263–284.

Stock, J. H. and Watson, M. W. (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, 20(2):147–162.

Vicente, M. R., López-Menéndez, A. J., and Pérez, R. (2015). Forecasting unemployment with internet search data: Does it help to improve predictions when job destruction is skyrocketing? *Technological Forecasting and Social Change*, 92:132–139.

Vosen, S. and Schmidt, T. (2011). Forecasting private consumption: survey-based indicators vs. Google Trends. *Journal of Forecasting*, 30(6):565–578.

Welch, I. and Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies*, 21(4):1455–1508.

Wu, L. and Brynjolfsson, E. (2015). The future of prediction: How Google searches foreshadow housing prices and sales. In *Economic analysis of the digital economy*, pages 89–118. University of Chicago Press.

Yang, X., Pan, B., Evans, J. A., and Lv, B. (2015). Forecasting Chinese tourist volume with search engine data. *Tourism Management*, 46:386–397.

A. Appendix

A.1. Descriptive statistics

Table 4: Descriptive statistics.

	Observations	*Mean*	*Standard deviation*	*Skewness*	*Kurtosis*
Unemployment rate	168	6.40	1.84	0.62	-1.13
Initial claims	731	355.36	88.64	1.25	1.46
Google Trends index	731	27.17	15.43	1.15	0.74

Notes: This table shows the descriptive statistics of the seasonally adjusted US unemployment rate in monthly frequency, the seasonally adjusted number of initial claims in weekly frequency, and the raw Google Trends index for the keyword “unemployment” in weekly frequency. The samples start in January 2004 and end in December 2017.

A.2. FRED-MD database

The TCODE column denotes the following data transformation for a series x: (1) no transformation; (2) Δx_t; (3) $\Delta^2 x_t$; (4) $log(x_t)$; (5) $\Delta log(x_t)$; (6) $\Delta^2 log(x_t)$; (7) $\Delta(x_t/x_{t-1} - 1.0)$. The FRED column gives mnemonics in FRED followed by a short description.

Some series require adjustments to the raw data available in FRED. These variables are tagged by an asterisk to indicate that they have been adjusted and thus differ from the series from the source. For a detailed summary of the adjustments see McCracken and Ng (2016).

Group 1. Output and income

	ID	*tcode*	*FRED*	*Description*
1	1	5	RPI	Real Personal Income
2	2	5	W875RX1	Real personal income ex transfer receipts
3	6	5	INDPRO	IP Index
4	7	5	IPFPNSS	IP: Financial Products and Nonindustrial Supplies
5	8	5	IPFINAL	IP: Final Products (Market Group)
6	9	5	IPCONGD	IP: Consumer Goods
7	10	5	IPDCONGD	IP: Durable Consumer Goods
8	11	5	IPNCONGD	IP: Nondurable Consumer Goods
9	12	5	IPBUSEQ	IP: Business Equipment
10	13	5	IPMAT	IP: Materials
11	14	5	IPDMAT	IP: Durable Materials
12	15	5	IPNMAT	IP: Nondurable Materials
13	16	5	IPMANSICS	IP: Manufacturing (SIC)
14	17	5	IPB51222s	IP: Residential Utilities
15	18	5	IPFUELS	IP: Fuels
16	19	1	NAPMPI	ISM Manufacturing: Production Index
17	20	2	CUMFNS	Capacity Utilization: Manufacturing

Group 2: Labor market

	ID	*tcode*	*FRED*	*Description*
1	21*	2	HWI	Help-Wanted Index for United States
2	22*	2	HWIURATIO	Ratio of Help Wanted/No. Unemployed
3	23	5	CLF160OV	Civilian Labor Force
4	24	5	CE160V	Civilian Employment
5	25	2	UNRATE	Civilian Unemployment Rate
6	26	2	UEMPMEAN	Average Duration of Unemployment (Weeks)
7	27	5	UEMPLT5	Civilians Unemployed - Less Than 5 Weeks
8	28	5	UEMP5TO14	Civilians Unemployed for 5-14 Weeks
9	29	5	UEMP15OV	Civilians Unemployed - 15 Weeks and Over
10	30	5	UEMP15T26	Civilians Unemployed for 15-26 Weeks
11	31	5	UEMP27OV	Civilians Unemployed for 27 Weeks and Over
12	32*	5	CLAIMSx	Initial Claims
13	33	5	PAYEMS	All Employees: Total nonfarm
14	34	5	USGOOD	All Employees: Goods-Producing Industries
15	35	5	CES1021000001	All Employees: Mining and Logging: Industries
16	36	5	USCONS	All Employees: Construction
17	37	5	MANEMP	All Employees: Manufacturing
18	38	5	DMANEMP	All Employees: Durable Goods
19	39	5	NDMANEMP	All Employees: Nondurable Goods
20	40	5	SRVPRD	All Employees: Service-Providing Industries
21	41	5	USTPU	All Employees: Trade, Transportation and Utilities
22	42	5	USWTRADE	All Employees: Wholesale Trade
23	43	5	USTRADE	All Employees: Retail Trade
24	44	5	USFIRE	All Employees: Financial Activities
25	45	5	USGOVT	All Employees: Government
26	46	1	CES0600000007	Avg Weekly Hours: Goods-Producing
27	47	2	AWOTMAN	Avg Weekly Overtime Hours: Manufacturing
28	48	1	AWHMAN	Avg Weekly Hours: Manufacturing
29	49	1	NAPMEI	ISM Manufacturing: Employment Index
30	127	6	CES0600000008	Avg Hourly Earnings: Goods-Producing
31	128	6	CES2000000008	Avg Hourly Earnings: Construction
32	129	6	CES3000000008	Avg Hourly Earnings: Manufacturing

Group 3: Housing

	ID	*tcode*	*FRED*	*Description*
1	50	4	HOUST	Housing Starts: Total New Privately Owned
2	51	4	HOUSTNE	Housing Starts: Northeast
3	52	4	HOUSTMW	Housing Starts: Midwest
4	53	4	HOUSTS	Housing Starts: South
5	54	4	HOUSTW	Housing Starts: West
6	55	4	PERMIT	New Private Housing Permits (SAAR)
7	56	4	PERMITNE	New Private Housing Permits: Northeast (SAAR)
8	57	4	PERMITMW	New Private Housing Permits: Midwest (SAAR)
9	58	4	PERMITS	New Private Housing Permits: South (SAAR)
10	59	4	PERMITW	New Private Housing Permits: West (SAAR)

Group 4: Consumption, orders and inventories

	ID	*tcode*	*FRED*	*Description*
1	3	5	DPCERA3M086SBEA	Real personal consumption expenditures
2	4*	5	CMRMTSPLx	Real Manu. and Trade Industries Sales
3	5*	5	RETAILx	Retail and Food Services Sales
4	60	1	NAPM	ISM: PMI Composite Index
5	61	1	NAPMNOI	ISM: New Orders Index
6	62	1	NAPMSDI	ISM: Supplier Deliveries Index
7	63	1	NAPMII	ISM: Inventories Index
8	64	5	ACOGNO	New Orders for Consumer Goods
9	65*	5	AMDMNOx	New Orders for Durable Goods
10	66*	5	ANDENOx	New Orders for Nondefense Capital Goods
11	67*	5	AMDMUOx	Unfilled Orders for Durable Goods
12	68*	5	BUSINVx	Total Business Inventories
13	69*	2	ISRATIOx	Total Business: Inventories to Sales Ratio
14	130*	2	UMSCENTx	Consumer Sentiment Index

Group 5: Money and credit

	ID	*tcode*	*FRED*	*Description*
1	70	6	M1SL	M1 Money Stock
2	71	6	M2SL	M2 Money Stock
3	72	5	M2REAL	Real M2 Money Stock
4	73	6	AMBSL	St. Louis Adjusted Monetary Base
5	74	6	TOTRESNS	Total Reserves of Depository Institutions
6	75	7	NONBORRES	Reserves of Depository Institutions
7	76	6	BUSLOANS	Commercial and Industrial Loans
8	77	6	REALLN	Real Estate Loans at All Commercial Banks
9	78	6	NONREVSL	Total Nonrevolving Credit
10	79*	2	CONSPI	Nonrevolving consumer credit to Personal Income
11	131	6	MZMSL	MZM Money Stock
12	132	6	DTCOLNVHFNM	Consumer Motor Vehicle Loans Outstanding
13	133	6	DTCTHFNM	Total Consumer Loans and Leases Outstanding
14	134	6	INVEST	Securities in Bank Credit at All Commercial Banks

Group 6: Interest and exchange rates

	ID	*tcode*	*FRED*	*Description*
1	84	2	FEDFUNDS	Effective Federal Funds Rate
2	85*	2	CP3Mx	3-Month AA Financial Commercial Paper Rate
3	86	2	TB3MS	3-Month Treasury Bill
4	87	2	TB6MS	6-Month Treasury Bill
5	88	2	GS1	1-Year Treasury Rate
6	89	2	GS5	5-Year Treasury Rate
7	90	2	GS10	10-Year Treasury Rate
8	91	2	AAA	Moody's Seasoned Aaa Corporate Bond Yield
9	92	2	BAA	Moody's Seasoned Baa Corporate Bond Yield
10	93*	1	COMPAPFFx	3-Month Commercial Paper Minus FEDFUNDS
11	94	1	TB3SMFFM	3-Month Treasury C Minus FEDFUNDS
12	95	1	TB6SMFFM	6-Month Treasury C Minus FEDFUNDS
13	96	1	T1YFFM	1-Year Treasury C Minus FEDFUNDS
14	97	1	T5YFFM	5-Year Treasury C Minus FEDFUNDS
15	98	1	T10YFFM	10-Year Treasury C Minus FEDFUNDS
16	99	1	AAAFFM	Moody's Aaa Corporate Bond Minus FEDFUNDS
17	100	1	BAAFFM	Moody's Baa Corporate Bond Minus FEDFUNDS
18	101	5	TWEXMMTH	Trade Weighted U.S. Dollar Index: Major Currencies
19	102*	5	EXSZUSx	Switzerland / U.S. Foreign Exchange Rate
20	103*	5	EXJPUSx	Japan / U.S. Foreign Exchange Rate
21	104*	5	EXUSUKx	U.S. / U.K. Foreign Exchange Rate
22	105*	5	EXCAUSx	Canada / U.S. Foreign Exchange Rate

Group 7: Prices

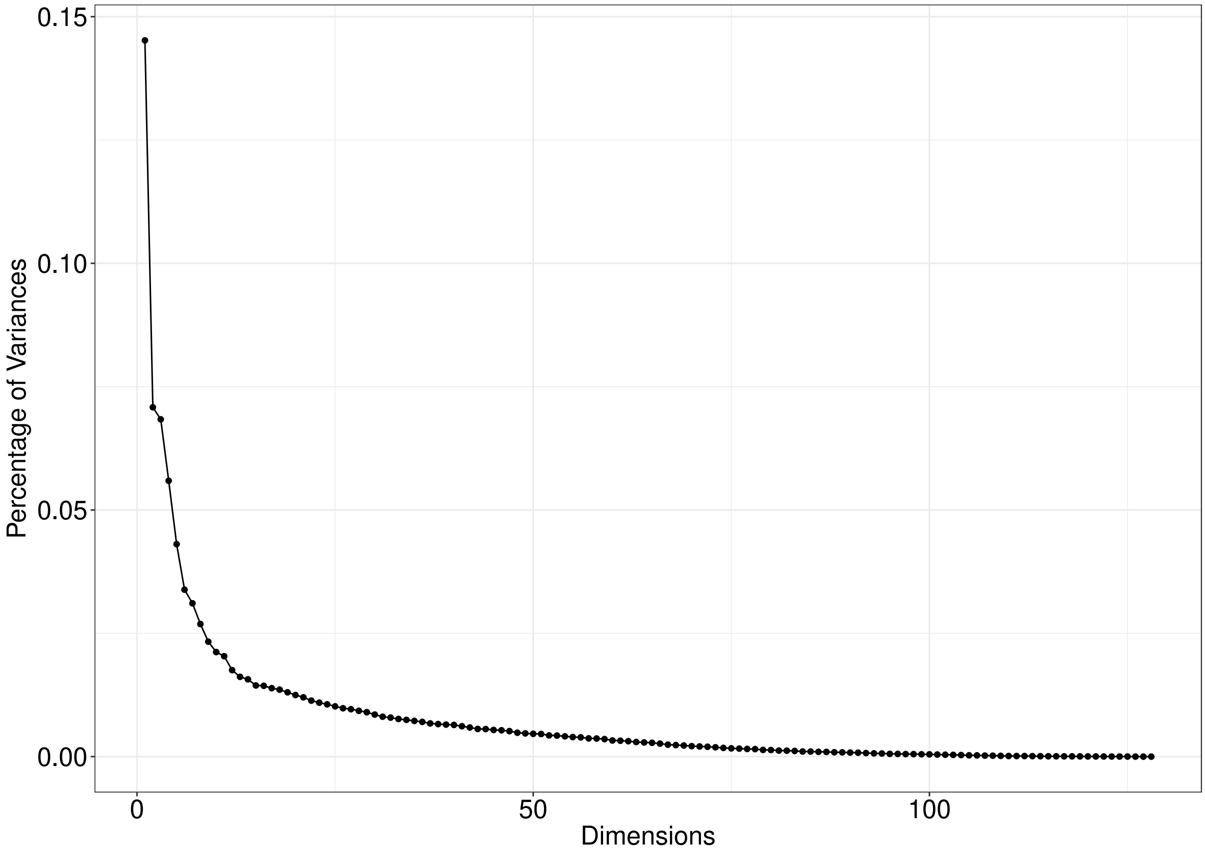
	ID	*tcode*	*FRED*	*Description*
1	106	6	WPSFD49207	PPI: Finished Goods
2	107	6	WPSFD49502	PPI: Finished Consumer Goods
3	108	6	WPSID61	PPI: Intermediate Materials
4	109	6	WPSID62	PPI: Crude Materials
5	110*	6	OILPRICEx	Crude Oil, spliced WTI and Cushing
6	111	6	PPICMM	PPI: Metals and metal products
7	112	1	NAPMPRI	ISM Manufacturing: Prices Index
8	113	6	CPIAUCSL	CPI: All Items
9	114	6	CPIAPPSL	CPI: Apparel
10	115	6	CPITRNSL	CPI: Transportation
11	116	6	CPIMEDSL	CPI: Medical Care
12	117	6	CUSR0000SAC	CPI: Commodities
13	118	6	CUSR0000SAD	CPI: Durables
14	119	6	CUSR0000SAS	CPI: Service
15	120	6	CPIULFSL	CPI: All Items less Food
16	121	6	CUSR0000SA0L2	CPI: All Items less Shelter
17	122	6	CUSR0000SA0L5	CPI: All Items less Medical Care
18	123	6	PCEPI	Personal Cons. Expend.: Chain Index
19	124	6	DDURRG3M086SBEA	Personal Cons. Expend.: Durable Goods
20	125	6	DNDGRG3M086SBEA	Personal Cons. Expend.: Nondurable Goods
21	126	6	DSERRG3M086SBEA	Personal Cons. Expend.: Services

Group 8: Stock market

	ID	*tcode*	*FRED*	*Description*
1	80*	5	S&P 500	S&P's Common Stock Price Index: Composite
2	81*	5	S&P: indust	S&P's Common Stock Price Index: Industrials
3	82*	2	S&P div yield	S&P's Composite Common Stock: Dividend Yield
4	83*	5	S&P PE ratio	S&P's Composite Common Stock: Price-Earnings Ratio

A.3. Scree plot of the estimated factors

Figure 4: Scree plot of estimated factors from the FRED-MD database.



Notes: The first four estimated factors explain 34 % of the variance of the sample. The PC_{p2} information criterion suggested by Bai and Ng (2002) proposes eight factors. Eight factors explain 47.5 % of the variance. Table 5 shows the corresponding values of this figure.

Table 5: Explained variance of the estimated factors from FRED-MD database.

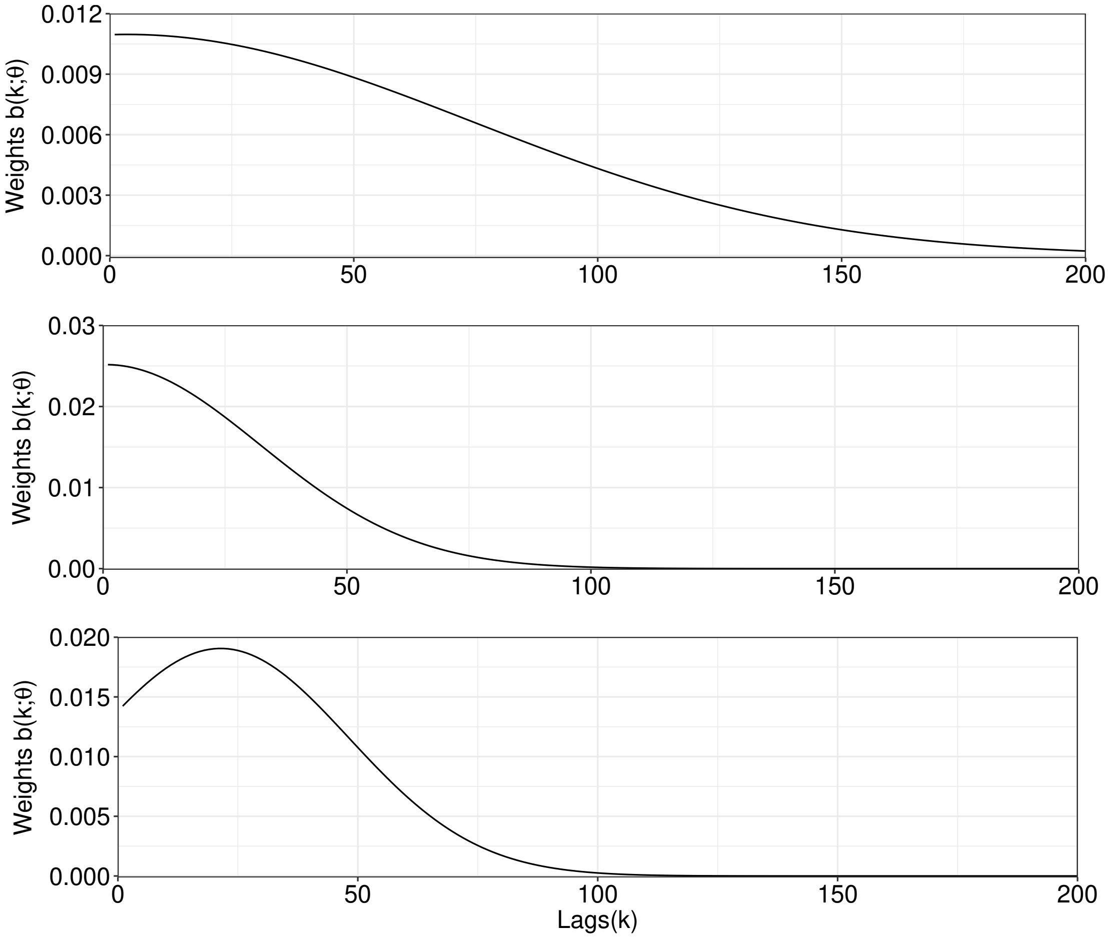
Dimension	*Variance*	*Cumulative variance*	*Dimension*	*Variance*	*Cumulative variance*
1	14.51949	14.51949	65	0.27976	95.18279
2	7.08187	21.60136	66	0.26382	95.44661
3	6.83728	28.43864	67	0.24229	95.68890
4	5.59214	34.03078	68	0.23373	95.92262
5	4.30795	38.33873	69	0.22531	96.14794
6	3.38398	41.72270	70	0.21244	96.36038
7	3.10758	44.83028	71	0.20594	96.56631
8	2.68821	47.51849	72	0.20020	96.76652
9	2.32999	49.84848	73	0.18902	96.95554
10	2.11969	51.96817	74	0.17707	97.13261
11	2.03620	54.00437	75	0.16772	97.30032
12	1.75344	55.75781	76	0.16269	97.46302
13	1.61722	57.37503	77	0.15540	97.61842
14	1.56662	58.94165	78	0.15239	97.77081
15	1.44283	60.38449	79	0.13570	97.90651
16	1.43543	61.81992	80	0.13419	98.04070
17	1.38754	63.20746	81	0.12289	98.16359
18	1.35770	64.56516	82	0.12103	98.28462
19	1.30348	65.86864	83	0.11503	98.39966
20	1.24842	67.11707	84	0.10369	98.50334
21	1.20290	68.31996	85	0.10270	98.60605
22	1.13536	69.45532	86	0.09745	98.70350
23	1.09385	70.54917	87	0.09537	98.79887
24	1.06028	71.60945	88	0.08827	98.88713
25	1.01918	72.62863	89	0.08642	98.97355
26	0.98102	73.60965	90	0.08236	99.05591
27	0.96153	74.57118	91	0.07944	99.13535
28	0.92869	75.49987	92	0.07267	99.20802
29	0.90057	76.40044	93	0.06616	99.27417
30	0.85245	77.25289	94	0.06366	99.33783
31	0.80921	78.06210	95	0.05574	99.39357
32	0.79091	78.85301	96	0.05448	99.44805
33	0.76288	79.61589	97	0.05199	99.50004
34	0.74521	80.36111	98	0.05098	99.55102
35	0.72364	81.08474	99	0.04735	99.59837
36	0.70437	81.78911	100	0.04593	99.64430
37	0.67319	82.46230	101	0.04189	99.68618
38	0.65986	83.12216	102	0.03859	99.72477
39	0.64926	83.77142	103	0.03530	99.76008

(Continued on next page)

Dimension	*Variance*	*Cumulative variance*	*Dimension*	*Variance*	*Cumulative variance*
40	0.64206	84.41348	104	0.03213	99.79220
41	0.61635	85.02983	105	0.02697	99.81917
42	0.59068	85.62050	106	0.02501	99.84418
43	0.56112	86.18162	107	0.02110	99.86528
44	0.55857	86.74019	108	0.01885	99.88413
45	0.54078	87.28097	109	0.01565	99.89978
46	0.53385	87.81481	110	0.01245	99.91223
47	0.51686	88.33168	111	0.01186	99.92409
48	0.48480	88.81647	112	0.01155	99.93564
49	0.47084	89.28731	113	0.00967	99.94532
50	0.46263	89.74994	114	0.00857	99.95388
51	0.45889	90.20882	115	0.00725	99.96114
52	0.43028	90.63911	116	0.00687	99.96800
53	0.42648	91.06559	117	0.00562	99.97362
54	0.41121	91.47680	118	0.00518	99.97880
55	0.39707	91.87388	119	0.00492	99.98372
56	0.39053	92.26440	120	0.00374	99.98746
57	0.36999	92.63440	121	0.00330	99.99076
58	0.36733	93.00172	122	0.00264	99.99339
59	0.35549	93.35721	123	0.00212	99.99552
60	0.32767	93.68488	124	0.00185	99.99737
61	0.32207	94.00695	125	0.00170	99.99907
62	0.31250	94.31944	126	0.00067	99.99974
63	0.29613	94.61557	127	0.00021	99.99995
64	0.28745	94.90303	128	0.00005	100.00000

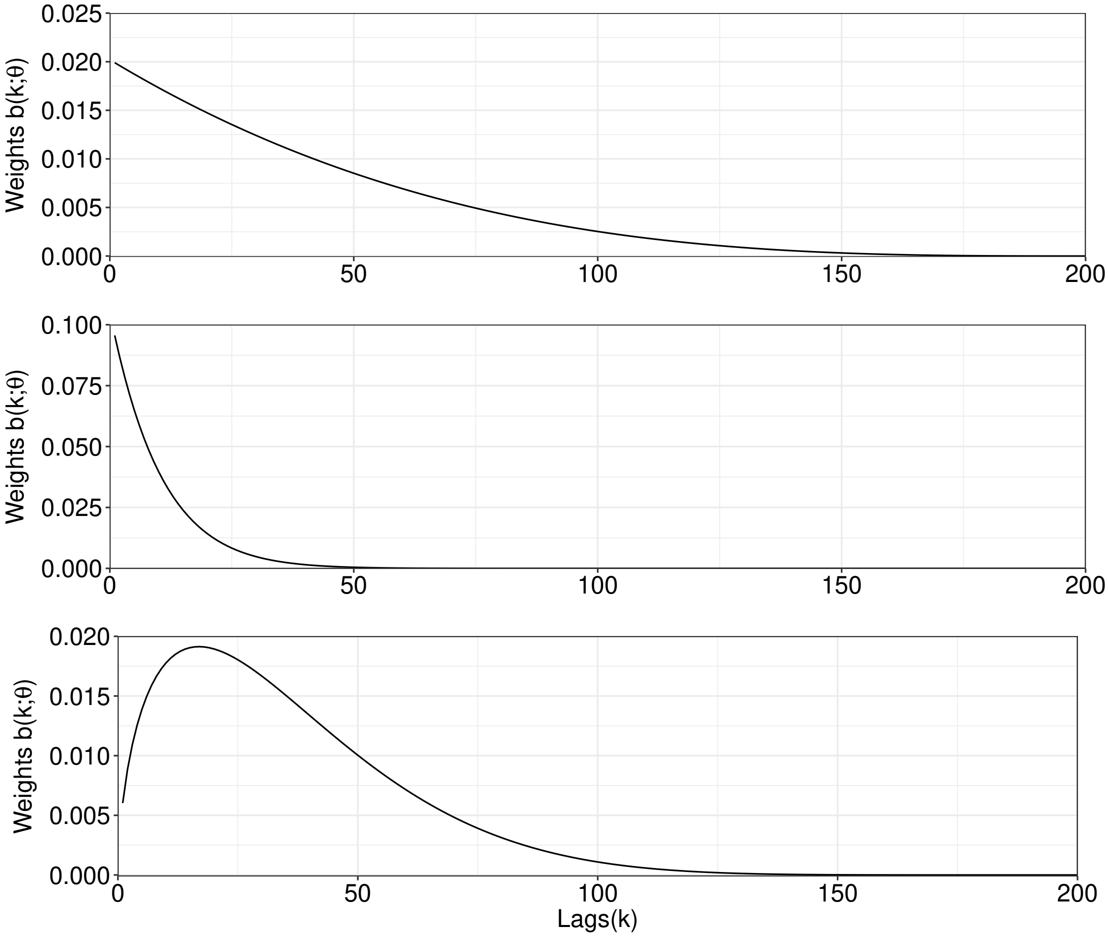
A.4. MIDAS specification weights

Figure 5: Exponential Almon polynomial MIDAS weights.



Notes: The figure shows three different shapes of the exponential Almon specification stated in (3.7). The shapes are determined by the values of the parameters $\theta = [\theta_1;\theta_2]$. The top panel shows slowly declining weights ($\theta_1 = 0.007$ and $\theta_2 = -0.0001$). The middle panel shows rapidly declining weights ($\theta_1 = 0.006$ and $\theta_2 = -0.0005$). The bottom panel shows weights that are hump-shaped ($\theta_1 = 0.03$ and $\theta_2 = -0.0007$). The values of θ_1, θ_2 and k are chosen only to illustrate the flexibility.

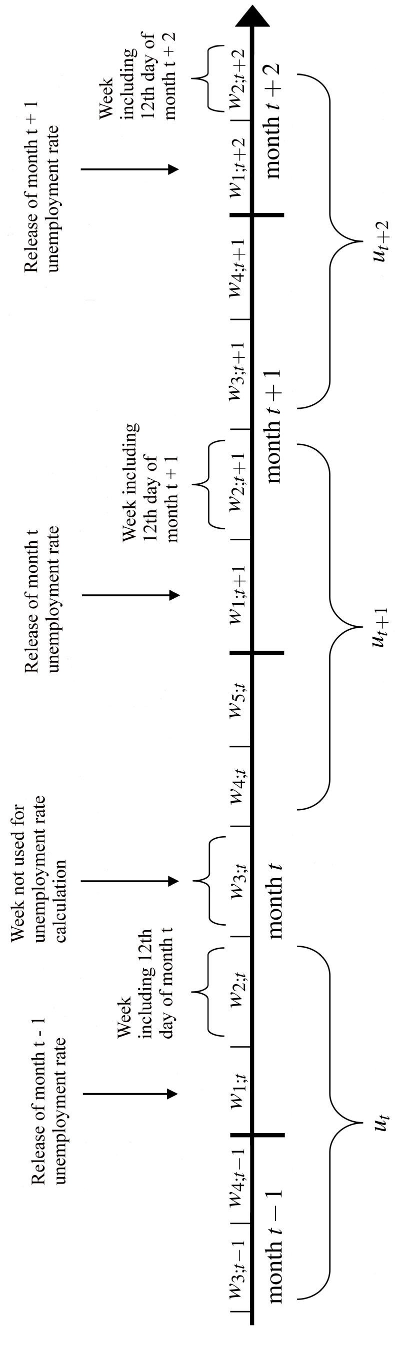
Figure 6: Beta polynomial MIDAS weights.



Notes: The figure shows three different shapes of the Beta specification stated in (3.8). The shapes are determined by the values of the parameters $\theta = [\theta_1; \theta_2]$. The top panel shows slowly declining weights ($\theta_1 = 1$ and $\theta_2 = 4$). The middle panel shows rapidly declining weights ($\theta_1 = 1$ and $\theta_2 = 20$). The bottom panel shows weights that are hump-shaped ($\theta_1 = 1.6$ and $\theta_2 = 7.5$). The values of θ_1, θ_2 and k are chosen only to illustrate the flexibility.

A.5. Timing of the US unemployment rate calculation

Figure 7: Timing of the US unemployment rate calculation.



Notes: The figure shows the exact timing of the US unemployment rate calculation as described in D'Amuri and Marcucci (2017).

A.6. Monthly relative RMSFE results compared to the D'Amuri and Marcucci (2017)-benchmark

A.6.1. Google Trends index

Table 6: Google Trends index: DI-AR, lag with Beta weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
January	1.083	1.008	1.000	0.989	0.910*	0.903**	0.940*	0.924	0.976	1.048	1.178	1.079
February	0.685	0.803	0.874	0.940	1.271	1.017	0.954	1.005	0.992	0.987	1.098	1.273
March	1.047	1.317	0.902	1.015	1.025	1.021	0.941	0.940	1.047	0.936	1.072	1.040
April	0.766**	1.127	1.172	1.014	0.957*	0.957	1.125	0.957	0.996	1.070	0.933	0.907*
May	0.863	1.257	1.400	1.179	1.163	0.980	0.914	1.043	0.972	0.997	1.027	1.005
June	1.279	1.017	1.338	1.387	0.963	1.090	1.138	1.034	1.096	0.984	0.948	1.011
July	1.429	0.928	0.975	1.267	1.035	0.764	1.059	1.009	0.930	1.132	1.009	0.916
August	0.725**	0.746	0.964	1.020	1.086	1.187	0.791	0.987	1.159	0.867	0.952	0.960
September	0.752	0.679*	0.858	0.934	1.090	0.806	1.078	0.878	1.062	1.107	0.938*	0.992
October	0.922	0.973	1.109	1.359	1.051	0.883*	0.931	0.936**	1.004	1.054	0.965	0.915
November	1.028	1.228	1.029	0.972	0.946	1.096	1.067	0.833	1.192	1.105	0.890	1.063
December	0.992	1.269	0.931	1.056	0.923*	0.822	0.937	0.944	0.794	1.122	1.150	0.986

Notes: The stars denote statistical significance at 10%(*), 5%(**) and 1%(***) level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

Table 7: Google Trends index: DI-AR with Beta weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
January	1.083	1.008	1.000	0.989	0.910*	0.903**	0.940*	0.924	0.976	1.048	1.178	1.079
February	0.685	0.803	0.874	0.940	1.271	1.017	0.954	1.005	0.992	0.987	1.098	1.273
March	1.047	1.317	0.902	1.015	1.025	1.021	0.941	0.940	1.047	0.936	1.072	1.040
April	0.766**	1.127	1.172	1.014	0.957*	0.957	1.125	0.957	0.996	1.070	0.933	0.907*
May	0.863	1.257	1.400	1.179	1.163	0.980	0.914	1.043	0.972	0.997	1.027	1.005
June	1.279	1.017	1.338	1.387	0.963	1.090	1.138	1.034	1.096	0.984	0.948	1.011
July	1.429	0.928	0.975	1.267	1.035	0.764	1.059	1.009	0.930	1.132	1.009	0.916
August	0.725**	0.746	0.964	1.020	1.086	1.187	0.791	0.987	1.159	0.867	0.952	0.960
September	0.752	0.679*	0.858	0.934	1.090	0.806	1.078	0.878	1.062	1.107	0.938*	0.992
October	0.922	0.973	1.109	1.359	1.051	0.883*	0.931	0.936**	1.004	1.054	0.965	0.915
November	1.028	1.228	1.029	0.972	0.946	1.096	1.067	0.833	1.192	1.105	0.890	1.063
December	0.992	1.269	0.931	1.056	0.923*	0.822	0.937	0.944	0.794	1.122	1.150	0.986

Notes: The stars denote statistical significance at 10%(*), 5%(**) and 1%(***) level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

Table 8: Google Trends index: DI with Beta weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
January	1.402	1.247	1.197	1.132	1.056	1.019	0.909**	0.986	0.945	1.049	1.316	1.233
February	1.046	1.214	1.103	1.037	1.229	1.003	1.048	0.998	1.016	0.966	1.110	1.270
March	0.976	1.520	1.327	1.222	1.310	1.006	0.963	0.963	0.994	1.008	0.946	1.093
April	0.657	1.401	1.393	1.215	1.191	1.039	1.097	1.016	1.034	1.025	0.973	0.785
May	0.934	1.225	1.543	1.276	1.600	1.075	1.030	1.024	0.972	1.048	1.133	1.001
June	1.128	1.092	1.362	1.444	1.520	1.206	1.173	1.096	1.076	1.007	1.039	1.049
July	1.225	0.809*	1.043	1.186	1.407	1.064	1.235	1.073	0.971	1.101	1.021	0.933
August	1.026	0.692	1.082	1.162	1.108	1.535	1.145	1.076	1.077	1.020	0.980	0.949
September	0.933	0.888	0.714*	0.851	0.986	0.803	1.442	1.091	1.108	1.128	1.009	1.028
October	1.118	1.127	1.201	1.148	1.137	0.951	0.924	1.222	1.152	1.143	1.011	0.979
November	0.978	1.276	1.075	1.100	1.002	1.002	1.056	0.851*	1.337	1.317	1.037	1.078
December	1.103	1.434	1.082	1.071	1.038	0.800	1.001	0.943	0.838	1.302	1.289	1.101

Notes: The stars denote statistical significance at 10%(*), 5%(**) and 1%(***) level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

Table 9: Google Trends index: DI-AR, lag with Nealmon weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
January	0.891	1.008	1.000	0.989	0.896**	0.904**	0.926**	0.969	0.981	1.000	1.138	1.063
February	0.612	0.698	0.855*	0.908	1.278	1.007	0.970	1.046	0.988	0.941	1.044	1.172
March	1.047	1.325	0.889	1.007	1.066	1.026	0.937	0.956	1.097	0.896	1.013	0.991
April	0.743**	1.127	1.130	0.993	1.026	0.957	1.113	0.952	1.013	1.067	0.882	0.788
May	0.874	1.181	1.192	1.086	1.191	0.966	0.923	1.039	0.987	1.045	0.994	0.914
June	1.225	1.017	1.185	1.078	0.967	1.042	1.124	1.023	1.083	0.996	0.962*	0.992
July	1.472	0.956	0.975	1.023	1.015	0.736	1.042	0.986	0.921	1.137	1.068	0.898
August	0.725**	0.752	1.000	1.024	0.945	1.221	0.835	0.984	1.095	0.878	0.976	0.967
September	0.808	0.620**	0.858	0.951	1.113	0.801	1.102	0.873	1.035	1.115	0.938	1.042
October	0.894	0.930	1.109	1.395	1.028	0.920*	0.898*	0.959	0.957	1.048	0.965	0.914
November	1.028	1.228	1.029	0.948	0.942	1.094	1.097	0.833	1.170	1.114	0.851	1.072
December	0.894	1.269	0.907	1.048	0.944	0.770	0.936	0.928	0.754	1.128	1.097	0.952

Notes: The stars denote statistical significance at $10\%(^{*})$, $5\%(^{**})$ and $1\%(^{***})$ level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

Table 10: Google Trends index: DI-AR with Nealmon weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
January	1.083	1.008	1.000	0.989	0.910*	0.903**	0.940*	0.924	0.976	1.048	1.178	1.079
February	0.685	0.803	0.874	0.940	1.271	1.017	0.954	1.005	0.992	0.987	1.098	1.273
March	1.047	1.317	0.902	1.015	1.025	1.021	0.941	0.940	1.047	0.936	1.072	1.040
April	0.766**	1.127	1.172	1.014	0.957*	0.957	1.125	0.957	0.996	1.070	0.933	0.907*
May	0.863	1.257	1.400	1.179	1.163	0.980	0.914	1.043	0.972	0.997	1.027	1.005
June	1.279	1.017	1.338	1.387	0.963	1.090	1.138	1.034	1.096	0.984	0.948	1.011
July	1.429	0.928	0.975	1.267	1.035	0.764	1.059	1.009	0.930	1.132	1.009	0.916
August	0.725**	0.746	0.964	1.020	1.086	1.187	0.791	0.987	1.159	0.867	0.952	0.960
September	0.752	0.679*	0.858	0.934	1.090	0.806	1.078	0.878	1.062	1.107	0.938*	0.992
October	0.922	0.973	1.109	1.359	1.051	0.883*	0.931	0.936**	1.004	1.054	0.965	0.915
November	1.028	1.228	1.029	0.972	0.946	1.096	1.067	0.833	1.192	1.105	0.890	1.063
December	0.992	1.269	0.931	1.056	0.923*	0.822	0.937	0.944	0.794	1.122	1.150	0.986

Notes: The stars denote statistical significance at $10\%(^{*})$, $5\%(^{**})$ and $1\%(^{***})$ level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

Table 11: Google Trends index: DI with Nealmon weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
January	1.402	1.247	1.197	1.132	1.056	1.019	0.909**	0.986	0.945	1.049	1.316	1.233
February	1.046	1.214	1.103	1.037	1.229	1.003	1.048	0.998	1.016	0.966	1.110	1.270
March	0.976	1.520	1.327	1.222	1.310	1.006	0.963	0.963	0.994	1.008	0.946	1.093
April	0.657	1.401	1.393	1.215	1.191	1.039	1.097	1.016	1.034	1.025	0.973	0.785
May	0.934	1.225	1.543	1.276	1.600	1.075	1.030	1.024	0.972	1.048	1.133	1.001
June	1.128	1.092	1.362	1.444	1.520	1.206	1.173	1.096	1.076	1.007	1.039	1.049
July	1.225	0.809*	1.043	1.186	1.407	1.064	1.235	1.073	0.971	1.101	1.021	0.933
August	1.026	0.692	1.082	1.162	1.108	1.535	1.145	1.076	1.077	1.020	0.980	0.949
September	0.933	0.888	0.714*	0.851	0.986	0.803	1.442	1.091	1.108	1.128	1.009	1.028
October	1.118	1.127	1.201	1.148	1.137	0.951	0.924	1.222	1.152	1.143	1.011	0.979
November	0.978	1.276	1.075	1.100	1.002	1.002	1.056	0.851*	1.337	1.317	1.037	1.078
December	1.103	1.434	1.082	1.071	1.038	0.800	1.001	0.943	0.838	1.302	1.289	1.101

Notes: The stars denote statistical significance at $10\%(^{*})$, $5\%(^{**})$ and $1\%(^{***})$ level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

A.6.2. Initial claims

Table 12: Initial claims: DI-AR, lag with Beta weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
January	1.130	1.173	0.994	1.002	0.873^{*}	0.926^{*}	0.951^{*}	0.878	0.934	1.093	1.173	1.116
February	0.433^{*}	0.725	0.938	0.932	1.277	0.966	0.960	1.048	0.972	0.939	1.104	1.224
March	0.951	1.325	0.872	1.065	1.049	1.014	0.933	0.944	1.043	0.832	0.967	1.035
April	0.754^{**}	1.019	1.119	0.972	1.087	0.952	1.116	0.950^{*}	1.013	1.040	0.861^{*}	0.777
May	0.863	1.136	1.217	1.156	1.204	1.008	0.915	1.045	0.991	0.988	1.038	0.881
June	1.552	1.060	1.314	1.223	1.076	0.983	1.162	1.005	1.098	0.996	0.948	0.996
July	1.291	0.891	0.965	1.196	1.010	0.747	1.052	1.045	0.920	1.134	1.029	0.960
August	0.649^{**}	0.740	0.986	1.020	1.089	1.266	0.787	1.023	1.116	0.881	0.965	0.949
September	0.722^{*}	0.588^{**}	0.915	0.884	0.993	0.840	1.078	0.930	1.040	1.124	0.946	1.005
October	0.671^{**}	1.174	1.016	1.387	0.827	0.903^{**}	0.942	0.978	1.008	1.023	0.986	0.933
November	0.870	1.066	1.005	0.915	1.006	1.031	1.040	0.877	1.168	1.109	0.896	1.076
December	1.155	1.054	0.896	1.074	0.946	0.816	0.950	0.916	0.808	1.176	1.094	1.014

Notes: The stars denote statistical significance at 10%(*), 5%(**) and 1%(***) level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

Table 13: Initial claims: DI-AR with Beta weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
January	1.130	1.173	0.994	1.002	0.873^{*}	0.926^{*}	0.951^{*}	0.878	0.934	1.093	1.173	1.116
February	0.433^{*}	0.725	0.938	0.932	1.277	0.966	0.960	1.048	0.972	0.939	1.104	1.224
March	0.951	1.325	0.872	1.065	1.049	1.014	0.933	0.944	1.043	0.832	0.967	1.035
April	0.754^{**}	1.019	1.119	0.972	1.087	0.952	1.116	0.950^{*}	1.013	1.040	0.861^{*}	0.777
May	0.863	1.136	1.217	1.156	1.204	1.008	0.915	1.045	0.991	0.988	1.038	0.881
June	1.552	1.060	1.314	1.223	1.076	0.983	1.162	1.005	1.098	0.996	0.948	0.996
July	1.291	0.891	0.965	1.196	1.010	0.747	1.052	1.045	0.920	1.134	1.029	0.960
August	0.649^{**}	0.740	0.986	1.020	1.089	1.266	0.787	1.023	1.116	0.881	0.965	0.949
September	0.722^{*}	0.588^{**}	0.915	0.884	0.993	0.840	1.078	0.930	1.040	1.124	0.946	1.005
October	0.671^{**}	1.174	1.016	1.387	0.827	0.903^{**}	0.942	0.978	1.008	1.023	0.986	0.933
November	0.870	1.066	1.005	0.915	1.006	1.031	1.040	0.877	1.168	1.109	0.896	1.076
December	1.155	1.054	0.896	1.074	0.946	0.816	0.950	0.916	0.808	1.176	1.094	1.014

Notes: The stars denote statistical significance at 10%(*), 5%(**) and 1%(***) level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

Table 14: Initial claims: DI with Beta weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
January	1.259	1.247	1.176	1.176	1.106	1.052	0.933^{**}	0.978	0.996	1.159	1.291	1.219
February	1.016	1.288	1.110	1.041	1.228	1.011	1.077	1.035	1.017	1.007	1.138	1.232
March	1.024	1.542	1.347	1.234	1.313	1.033	0.974	0.985	1.032	0.884	0.957	1.070
April	0.587^{*}	1.468	1.513	1.228	1.204	1.030	1.124	1.036	1.069	1.032	0.864^{*}	0.783
May	1.243	1.460	1.685	1.389	1.631	1.077	1.040	1.046	1.005	1.077	0.985	0.895
June	1.108	1.338	1.492	1.410	1.587	1.203	1.171	1.095	1.107	1.030	1.034	0.929^{*}
July	0.890	0.871	1.133	1.364	1.414	1.008	1.243	1.062	0.975	1.120	1.045	0.919
August	0.889	0.604	1.069	1.241	1.254	1.495	1.171	1.077	1.095	1.041	0.991	0.924
September	1.103	0.888	0.739^{**}	0.892	1.016	0.951	1.374	1.127	1.119	1.159	1.027	1.010
October	1.304	1.461	1.361	1.292	1.221	0.992	1.046	1.180	1.138	1.166	0.995	0.978
November	0.746	1.262	1.024	1.152	1.033	1.028	1.090	1.060	1.268	1.340	1.048	1.077
December	1.111	1.459	1.094	1.109	1.098	0.848	1.036	0.932	0.911	1.259	1.264	1.108

Notes: The stars denote statistical significance at 10%(*), 5%(**) and 1%(***) level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

Table 15: Initial claims: DI-AR, lag with Nealmon weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
January	1.017	0.988	0.904	0.975	0.896	0.927	0.959*	0.905	0.905*	1.062	1.186	1.086
February	0.586*	0.803	0.835*	0.963	1.248	1.028	0.960	1.032	0.987	0.894	1.094	1.174
March	0.724	1.238	0.996	1.097	1.062	1.028	0.938	0.938	1.044	0.915	0.973	1.039
April	0.743**	1.127	1.156	1.018	1.015	0.960	1.098	0.968	1.005	1.040	0.873*	0.819
May	0.786	1.170	1.497	1.126	1.142	1.010	0.941	1.020	0.978	1.007	1.045	0.924
June	1.414	1.043	1.129	1.209	1.026	1.069	1.159	1.045	1.066	0.997	0.961*	1.028
July	1.041	1.067	0.886	1.131	1.070	0.784	1.068	1.031	0.935	1.137	1.038	0.909
August	0.607**	0.659	1.082	0.922	1.069	1.249	0.848	0.985	1.047	0.896	0.943	0.942
September	0.956	0.707*	0.786*	1.098	0.991	0.832	1.107	0.916	1.079	1.084	0.963	0.993
October	0.922	1.241	0.878*	1.288	0.852	0.878**	0.937	0.942	1.055	1.076	0.986	0.923
November	0.941	1.221	0.990	1.059	0.929	0.988	1.057	0.888	1.133	1.112	0.891	1.061
December	1.291	1.130	0.887	1.124	0.925	0.839	0.945	0.926	0.889	1.152	1.123	1.013

Notes: The stars denote statistical significance at 10%(*), 5%(**) and 1%(***) level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

Table 16: Initial claims: DI-AR with Nealmon weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
January	1.189	1.269	0.960	1.037	0.890	0.911*	0.940*	0.891	0.934	1.102	1.055	1.100
February	0.661	0.851	0.889	0.867	1.276	0.992	0.963	1.025	0.975	0.947	1.064	1.179
March	1.000	1.256	0.802	1.037	1.093	1.008	0.943	0.949	1.050	0.874	0.951	0.976
April	0.777*	1.193	1.167	0.982	1.033	0.917	1.142	0.924	1.000	1.058	0.913	0.811
May	0.863	1.064	1.296	1.113	1.183	1.031	0.877	1.043	0.991	0.964	1.029	0.911
June	1.297	1.060	1.308	1.299	1.048	1.006	1.135	1.058	1.098	0.977	0.973	0.998
July	1.384	0.900	0.955	1.258	1.144	0.888	1.088	1.051	0.906	1.142	0.978	0.924
August	0.688**	0.740	1.007	1.004	1.066	1.147	0.793	0.979	1.125	0.889	0.957	0.949
September	0.909	0.721*	0.842*	0.987	0.998	0.843	1.078	0.942	1.012	1.086	0.949	1.008
October	0.866	1.090	1.064	1.371	0.926	0.906	0.919*	1.015	1.006	1.045	0.944	0.910
November	0.854	1.017	0.892	0.969	0.949	1.040	1.030	0.873	1.130	1.109	0.885	1.068
December	1.176	1.045	0.948	1.064	0.936	0.774	0.950	0.920	0.837	1.128	1.080	0.949

Notes: The stars denote statistical significance at 10%(*), 5%(**) and 1%(***) level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

Table 17: Initial claims: DI with Nealmon weights.

	h=1	*h=2*	*h=3*	*h=4*	*h=5*	*h=6*	*h=7*	*h=8*	*h=9*	*h=10*	*h=11*	*h=12*
January	1.326	1.237	1.179	1.148	1.122	1.049	0.929*	0.982	0.988	1.181	1.180	1.175
February	1.132	1.181	1.060	0.983	1.224	0.994	1.079	1.025	1.027	1.026	1.117	1.191
March	0.900	1.633	1.288	1.192	1.315	1.026	0.980	1.001	1.027	0.965	0.951	1.063
April	0.682*	1.144	1.501	1.110	1.156	0.956	1.124	1.013	1.075	1.047	0.896	0.798
May	0.991	1.433	1.378	1.276	1.547	1.012	1.042	1.053	0.976	1.074	0.985	0.858
June	1.168	1.364	1.420	1.372	1.612	1.119	1.148	1.097	1.097	1.032	1.032	0.902*
July	0.816	0.900	1.225	1.422	1.328	1.018	1.189	1.036	0.968*	1.121	1.023	0.916
August	0.946	0.575	1.095	1.244	1.265	1.266	1.129	1.040	1.068	1.010	0.988	0.944*
September	1.043	0.888	0.751**	0.877	1.076	0.930	1.344	1.082	1.092	1.140	0.971**	1.026
October	1.245	1.241	1.391	1.309	1.208	0.963	1.066	1.117	1.099	1.139	0.989	0.968*
November	0.707	1.289	1.005	1.164	1.027	1.028	1.097	1.080	1.232	1.306	0.998	1.080
December	0.975	1.269	1.121	1.119	1.101	0.847	1.035	0.941	0.928	1.218	1.212	1.099

Notes: The stars denote statistical significance at 10%(*), 5%(**) and 1%(***) level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

A.6.3. Google Trends index and initial claims

Table 18: Google Trends index and initial claims: DI-AR, lag with Beta weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
January	1.287	1.046	1.003	1.000	0.901**	0.956	0.935**	0.896	0.997	1.069	1.180	1.095
February	0.661	0.881	0.832*	0.925	1.278	0.983	0.953	1.040	0.992	0.980	1.098	1.301
March	0.926	1.256	0.812	1.016	1.014	1.021	0.941	0.940	1.032	0.977	1.059	1.069
April	0.777**	1.144	1.162	0.979	1.021	0.957	1.120	0.957	0.977	1.062	0.957	0.862
May	0.934	1.181	1.456	1.191	1.163	0.982	0.911	1.043	0.975	0.986	1.082	1.012
June	1.523	1.051	1.238	1.266	1.013	0.991	1.141	1.022	1.093	0.981	0.948	1.003
July	1.384	0.851	1.000	1.271	1.153	0.705	1.058	1.041	0.930	1.136	1.004	0.903
August	0.688**	0.746	0.926	0.988	1.069	1.174	0.720	0.991	1.149	0.884	0.952	0.947
September	0.626**	0.604**	0.815*	1.000	1.078	0.822	1.078	0.889	1.073	1.111	0.927	0.985
October	0.975	1.013	1.064	1.342	1.049	0.877*	0.931	0.936**	0.987	1.027	0.962	0.909
November	1.021	1.235	1.014	0.955	0.953	1.125	1.055	0.833	1.161	1.007	0.890	1.061
December	0.966	1.269	0.962	1.048	0.957	0.821	0.969	0.917	0.783	1.092	1.137	1.022

Notes: The stars denote statistical significance at 10%(*), 5%(**) and 1%(***) level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

Table 19: Google Trends index and initial claims: DI-AR with Beta weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
January	1.287	1.046	1.003	1.000	0.901**	0.956	0.935**	0.896	0.997	1.069	1.180	1.095
February	0.661	0.881	0.832*	0.925	1.278	0.983	0.953	1.040	0.992	0.980	1.098	1.301
March	0.926	1.256	0.812	1.016	1.014	1.021	0.941	0.940	1.032	0.977	1.059	1.069
April	0.777**	1.144	1.162	0.979	1.021	0.957	1.120	0.957	0.977	1.062	0.957	0.862
May	0.934	1.181	1.456	1.191	1.163	0.982	0.911	1.043	0.975	0.986	1.082	1.012
June	1.523	1.051	1.238	1.266	1.013	0.991	1.141	1.022	1.093	0.981	0.948	1.003
July	1.384	0.851	1.000	1.271	1.153	0.705	1.058	1.041	0.930	1.136	1.004	0.903
August	0.688**	0.746	0.926	0.988	1.069	1.174	0.720	0.991	1.149	0.884	0.952	0.947
September	0.626**	0.604**	0.815*	1.000	1.078	0.822	1.078	0.889	1.073	1.111	0.927	0.985
October	0.975	1.013	1.064	1.342	1.049	0.877*	0.931	0.936**	0.987	1.027	0.962	0.909
November	1.021	1.235	1.014	0.955	0.953	1.125	1.055	0.833	1.161	1.007	0.890	1.061
December	0.966	1.269	0.962	1.048	0.957	0.821	0.969	0.917	0.783	1.092	1.137	1.022

Notes: The stars denote statistical significance at 10%(*), 5%(**) and 1%(***) level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

Table 20: Google Trends index and initial claims: DI with Beta weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	*h=11*	*h=12*
January	1.339	1.169	1.179	1.132	1.041	1.025	0.898**	0.987	0.958	1.049	1.392	1.249
February	0.968	1.323	1.078	1.048	1.186	1.003	1.043	1.009	1.016	1.004	1.110	1.307
March	0.787	1.461	1.329	1.219	1.302	1.006	0.976	0.957	1.001	1.005	0.953	1.134
April	0.743	1.240	1.358	1.219	1.197	1.041	1.100	1.014	1.039	1.029	0.949	0.763
May	0.915	1.308	1.510	1.349	1.611	1.077	1.041	1.033	0.983	1.055	1.119	1.003
June	1.128	1.092	1.356	1.364	1.610	1.206	1.177	1.096	1.084	1.007	1.017	1.084
July	1.155	0.809*	1.024	1.215	1.436	1.054	1.235	1.065	0.975	1.105	1.027	0.974
August	1.051	0.639	1.056	1.172	1.108	1.491	1.159	1.076	1.109	1.020	0.986	0.924
September	0.909	0.877	0.647*	0.884	0.982	0.804	1.459	1.094	1.116	1.170	1.039	1.002
October	1.140	1.139	1.173	1.119	1.142	0.934	0.925	1.228	1.164	1.144	1.008	0.974
November	0.949	1.372	1.043	1.079	0.985	1.002	1.046	0.851*	1.368	1.377	1.056	1.103
December	0.940	1.484	1.029	1.071	1.036	0.795	1.017	0.949	0.824	1.317	1.307	1.134

Notes: The stars denote statistical significance at 10%(*), 5%(**) and 1%(***) level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

Table 21: Google Trends index and initial claims: DI-AR, lag with Nealmon weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
January	0.947	0.992	0.969	0.985	0.860*	0.906**	0.942*	0.973	0.965	1.003	1.124	1.034
February	0.586	0.752	0.868	0.963	1.261	1.034	0.981	1.057	0.994	0.920	1.021	1.188
March	0.976	1.247	0.952	1.080	1.065	0.998	0.943	0.959	1.070	0.903	0.999	0.990
April	0.754**	1.240	1.125	1.027	0.973	0.907	1.111	0.968	1.000	1.060	0.849*	0.769
May	0.820	1.158	1.349	1.054	1.166	0.966	0.918	1.027	0.977	1.059	0.988	0.899
June	1.225	0.946	1.257	1.191	1.010	1.032	1.131	1.060	1.074	0.990	0.959**	1.001
July	1.541	1.083	1.098	1.146	1.030	0.693	1.031	0.966	0.925	1.133	1.094	0.895
August	0.607**	0.716	1.082	0.926	0.971	1.212	0.848	0.998	1.038	0.867	0.955**	0.961
September	0.752	0.747*	0.757*	1.235	1.078	0.766	1.083	0.878	1.035	1.064	0.938*	1.037
October	1.072	1.053	0.932	1.280	0.972	0.902*	0.901	0.923**	0.984	1.064	0.964	0.902
November	1.028	1.200	1.038	0.979	0.880**	1.039	1.116	0.815	1.193	1.067	0.886	1.062
December	0.940	1.262	0.913	1.116	0.949	0.832	0.954	0.965	0.796	1.035	1.115	0.976

Notes: The stars denote statistical significance at $10\%(^{*})$, $5\%(^{**})$ and $1\%(^{***})$ level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

Table 22: Google Trends index and initial claims: DI-AR with Nealmon weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
January	1.339	1.108	0.937	1.005	0.863**	0.938	0.933**	0.895	0.999	1.075	1.096	1.105
February	0.750	0.960	0.819	0.891	1.269	0.969	0.963	1.012	0.989	1.017	1.113	1.274
March	1.024	1.125	0.858	1.018	1.049	1.015	0.934	0.940	1.042	1.006	1.097	1.022
April	0.754**	1.144	1.146	0.937	1.007	0.915	1.160	0.931	1.000	1.078	0.981	0.942
May	0.842	1.225	1.334	1.118	1.135	1.016	0.871	1.045	0.972	0.966	1.053	0.962
June	1.348	1.043	1.244	1.485	0.997	1.018	1.146	1.072	1.087	0.967	0.948	0.955
July	1.486	0.974	0.975	1.285	1.131	0.780	1.060	1.000	0.912	1.132	0.971	0.939
August	0.725*	0.746	1.007	1.004	1.062	1.174	0.744	0.990	1.146	0.868	0.945	0.943
September	0.885	0.665**	0.786**	0.964	1.088	0.776	1.100	0.891	1.065	1.071	0.914	0.995
October	1.025	0.900	1.131	1.322	1.051	0.891	0.937	0.982	1.024	1.063	0.953	0.928
November	0.986	1.207	1.000	0.969	0.911	1.091	1.066	0.884	1.179	1.037	0.874	1.058
December	0.957	1.312	0.995	1.024	0.934	0.792	0.957	0.945	0.834	1.089	1.080	0.927*

Notes: The stars denote statistical significance at $10\%(^{*})$, $5\%(^{**})$ and $1\%(^{***})$ level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

Table 23: Google Trends index and initial claims: DI with Nealmon weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
January	1.565	1.272	1.140	1.142	1.160	1.057	0.951*	0.994	1.041	1.194	1.305	1.162
February	1.016	1.318	1.017	1.012	1.231	1.009	1.078	1.020	1.036	1.068	1.149	1.294
March	0.976	1.468	1.258	1.173	1.351	1.024	0.996	0.978	1.032	0.953	0.976	1.078
April	0.809	1.641	1.464	1.108	1.195	0.983	1.124	1.011	1.071	1.031	0.940	0.831
May	1.000	1.598	1.606	1.267	1.532	1.039	1.030	1.031	0.967	1.040	0.928*	0.948
June	1.279	1.370	1.481	1.503	1.517	1.109	1.128	1.115	1.081	1.019	1.025	0.935
July	0.957	0.919	1.229	1.402	1.457	1.014	1.172	1.027	0.975	1.143	1.014	0.939
August	0.973	0.672	1.069	1.315	1.233	1.396	1.129	1.052	1.050	1.012	0.971	0.940*
September	1.022	1.038	0.757	0.947*	1.103	0.924	1.431	1.074	1.069	1.133	0.976	1.020
October	1.342	1.284	1.367	1.167	1.314	0.992	1.052	1.101	1.100	1.092	0.988	0.961*
November	0.697	1.365	0.970	1.100	0.962	1.112	1.138	1.068	1.242	1.317	0.992	1.057
December	1.041	1.269	1.111	1.140	1.052	0.812	1.067	0.947	0.943	1.246	1.204	1.036

Notes: The stars denote statistical significance at $10\%(^{*})$, $5\%(^{**})$ and $1\%(^{***})$ level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

A.6.4. Factors and Google Trends index

Table 24: Factors and Google Trends index: DI-AR, lag with Beta weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
January	1.619	1.584	1.065	1.055	1.194	1.117	1.033	0.902	0.916	1.002	1.189	0.874
February	1.186	0.659	0.889*	1.188	1.350	1.180	0.945	0.938	1.003	0.857	0.911	1.340
March	2.070	0.816	0.885	1.447	1.603	1.066	1.476	1.002	1.087	0.800	1.522	0.961
April	1.548	4.029	0.836*	1.421	1.067	0.970	1.195	1.029	1.020	1.057	1.199	1.344
May	2.604	1.792	1.631	1.215	0.924	1.325	1.173	1.041	1.076	1.065	1.421	1.071
June	1.919	1.395	1.178	1.477	1.085	1.015	1.149	1.126	1.151	1.022	0.934	1.031
July	1.620	1.253	1.973	1.329	1.425	0.664	1.173	1.316	0.995	1.155	1.043	1.623
August	3.449	0.885	1.558	1.648	1.376	1.666	1.279	0.574	1.335	0.972	0.948	0.886**
September	3.086	1.083	1.286	1.182	1.275	0.664*	1.488	0.838	1.386	0.991	1.023	0.983
October	1.378	2.156	1.765	1.583	1.821	0.967	1.119	1.133	0.937	0.885	1.320	1.071
November	1.424	2.042	1.431	1.489	1.839	1.425	1.475	1.078	1.481	1.149	0.614	1.302
December	1.623	1.958	0.896	1.554	1.212	0.679	0.946	0.789**	1.146	1.048	0.691***	1.036

Notes: The stars denote statistical significance at 10%(*), 5%(**) and 1%(***) level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

Table 25: Factors and Google Trends index: DI-AR with Beta weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
January	1.203	1.262	1.035	1.231	1.060	1.208	0.957	0.924	1.067	1.147	1.222	0.966
February	0.661	0.628	0.981	1.151	1.304	1.125	1.044	0.959	0.990	1.139	1.156	1.361
March	1.773	1.274	0.659	1.352	1.488	1.042	1.107	1.029	1.041	0.826	1.305	1.109
April	0.657	1.754	1.012	1.249	1.059	1.016	1.211	0.977	1.043	1.040	0.930	1.085
May	1.198	1.784	2.049	1.110	0.841	1.065	1.089	0.993	1.003	1.122	0.947	1.081
June	1.610	1.147	1.295	1.545	1.016	1.319	1.249	1.135	1.165	0.969	0.983	1.039
July	2.716	1.182	1.738	1.206	1.050	0.709	1.046	1.105	0.942	1.114	1.013	1.208
August	2.575	1.322	1.082	1.394	1.156	1.457	0.879	0.619	1.211	0.899	0.977	0.957
September	2.236	0.635*	0.899	0.902	1.065	0.748	1.295	0.776	1.049	1.034	0.847	1.001
October	1.817	1.831	1.689	1.505	1.166	0.883	1.017	1.023	0.889*	0.925*	1.016	0.915
November	1.242	1.722	1.358	1.226	1.074	1.184	1.133	0.752	1.287	1.000	0.649	1.148
December	1.420	2.046	0.532	1.450	1.231	0.737*	0.893	0.842	0.735	1.102	0.786***	0.982

Notes: The stars denote statistical significance at 10%(*), 5%(**) and 1%(***) level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

Table 26: Factors and Google Trends index: DI with Beta weights.

	h=1	*h=2*	*h=3*	*h=4*	*h=5*	*h=6*	*h=7*	*h=8*	*h=9*	*h=10*	*h=11*	*h=12*
January	1.034	1.431	1.356	1.287	1.331	1.157	0.926**	1.035	0.939	0.989	1.222	1.052
February	0.685	0.607	1.046	1.283	1.073	0.969	1.040	0.941	0.973	1.056	1.089	1.221
March	1.604	0.943	0.659	1.282	1.389	0.996	0.989	0.970	1.004	0.952	0.985	0.908
April	0.670	1.519	0.892	0.847	1.326	1.028	1.143	1.031	1.051	1.008	0.925	0.733*
May	1.112	1.433	1.871	0.842	1.032	1.106	1.081	0.933	0.934	1.055	1.066	0.963
June	1.462	0.955	1.164	1.610	0.946	0.983	1.287	1.104	1.133	0.993	1.041	1.097
July	2.236	0.956	1.125	1.410	1.678	0.701	0.976	1.189	1.027	1.135	1.040	1.158
August	1.556	1.038	1.082	1.309	1.174	1.669	0.783	0.638	1.050	0.951	0.952	0.888
September	2.449	0.635*	0.603*	0.793	1.193	0.701*	1.562	0.711	0.886**	1.169	0.955	1.040
October	1.500	2.422	1.558	1.267	1.241	0.985	0.958	1.114	0.912*	0.891	1.038	0.962
November	1.165	1.378	1.461	1.221	1.053	1.085	1.147	0.813	1.330	1.064	0.813	1.040
December	1.297	1.905	0.705	1.128	1.055	0.713*	0.961	0.857	0.808	1.187	0.896**	1.037

Notes: The stars denote statistical significance at 10%(*), 5%(**) and 1%(***) level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

Table 27: Factors and Google Trends index: DI-AR, lag with Nealmon weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
January	1.692	1.497	1.085	1.035	1.137	1.105	1.025	0.999	0.882	1.068	1.162	0.839^{**}
February	0.848	0.679	0.899	1.125	1.377	1.110	0.900	0.947	1.003	0.879	0.927	1.285
March	2.582	0.843	0.816	1.468	1.383	1.002	1.424	1.000	1.308	1.025	1.303	0.862
April	1.526	4.151	0.814^{**}	1.437	0.945	0.840	1.237	0.970	0.992	1.068	1.119	1.211
May	2.479	1.878	1.822	1.264	0.964	1.271	1.053	1.026	1.032	1.038	1.289	1.064
June	1.477	1.221	1.136	1.441	1.045	0.983	1.114	1.027	1.219	1.028	0.953	0.983
July	1.744	1.253	2.036	1.262	1.300	0.683	1.208	1.233	0.962	1.186	1.065	1.587
August	3.220	0.786	1.449	1.854	1.558	1.787	1.264	0.606	1.190	0.911	0.921^{**}	0.896^{*}
September	2.537	1.177	1.480	1.177	1.332	0.656^{*}	1.474	0.855	1.421	0.974	1.010	0.983
October	1.183	2.449	1.499	1.975	1.540	1.057	1.088	1.254	0.934^{*}	0.885	1.268	1.039
November	1.434	2.111	1.542	1.643	2.110	1.414	1.397	0.989	1.368	1.268	0.616	1.361
December	1.936	2.005	0.777	1.618	1.230	0.735	0.839	0.773^{**}	1.065	1.018	0.676^{***}	1.080

Notes: The stars denote statistical significance at 10%(*), 5%(**) and 1%(***) level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

Table 28: Factors and Google Trends index: DI-AR with Nealmon weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	*h=11*	*h=12*
January	1.203	1.262	1.035	1.231	1.060	1.208	0.957	0.924	1.067	1.147	1.222	0.966
February	0.661	0.628	0.981	1.151	1.304	1.125	1.044	0.959	0.990	1.139	1.156	1.361
March	1.773	1.274	0.659	1.352	1.488	1.042	1.107	1.029	1.041	0.826	1.305	1.109
April	0.657	1.754	1.012	1.249	1.059	1.016	1.211	0.977	1.043	1.040	0.930	1.085
May	1.198	1.784	2.049	1.110	0.841	1.065	1.089	0.993	1.003	1.122	0.947	1.081
June	1.610	1.147	1.295	1.545	1.016	1.319	1.249	1.135	1.165	0.969	0.983	1.039
July	2.716	1.182	1.738	1.206	1.050	0.709	1.046	1.105	0.942	1.114	1.013	1.208
August	2.575	1.322	1.082	1.394	1.156	1.457	0.879	0.619	1.211	0.899	0.977	0.957
September	2.236	0.635^{*}	0.899	0.902	1.065	0.748	1.295	0.776	1.049	1.034	0.847	1.001
October	1.817	1.831	1.689	1.505	1.166	0.883	1.017	1.023	0.889^{*}	0.925^{*}	1.016	0.915
November	1.242	1.722	1.358	1.226	1.074	1.184	1.133	0.752	1.287	1.000	0.649	1.148
December	1.420	2.046	0.532	1.450	1.231	0.737^{*}	0.893	0.842	0.735	1.102	0.786^{***}	0.982

Notes: The stars denote statistical significance at 10%(*), 5%(**) and 1%(***) level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

Table 29: Factors and Google Trends index: DI with Nealmon weights.

	h=1	*h=2*	*h=3*	*h=4*	*h=5*	*h=6*	*h=7*	*h=8*	*h=9*	*h=10*	*h=11*	*h=12*
January	1.034	1.471	1.356	1.287	1.323	1.153	0.921^{**}	1.029	0.911	1.032	1.250	1.061
February	0.685	0.607	1.046	1.275	1.063	0.961	1.040	0.935	0.967	1.031	1.126	1.233
March	1.604	0.943	0.659	1.282	1.381	0.983	0.984	0.968	1.000	0.937	0.939	0.920
April	0.682	1.506	0.871	0.845	1.314	1.024	1.139	1.026	1.051	1.007	0.925	0.704^{*}
May	1.168	1.367	1.833	0.842	1.032	1.086	1.071	0.922	0.929	1.055	1.047	0.902
June	1.462	1.051	1.100	1.573	0.987	0.983	1.268	1.096	1.121	0.987	1.039	1.059
July	2.273	1.000	1.180	1.307	1.643	0.732	0.976	1.172	1.021	1.120	1.035	1.144
August	1.606	1.038	1.102	1.342	1.140	1.669	0.812	0.655	1.032	0.944	0.939	0.880
September	2.441	0.635^{*}	0.572^{*}	0.793	1.193	0.701^{*}	1.571	0.733	0.890^{*}	1.136	0.941	1.029
October	1.565	2.427	1.558	1.267	1.249	0.998	0.968	1.124	0.927	0.891	1.014	0.951
November	1.108	1.334	1.438	1.245	1.035	1.085	1.147	0.824	1.333	1.074	0.814	1.011
December	1.366	1.910	0.705	1.121	1.053	0.710^{*}	0.954	0.847	0.851	1.207	0.912^{**}	1.037

Notes: The stars denote statistical significance at 10%(*), 5%(**) and 1%(***) level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

A.6.5. Factors and initial claims

Table 30: Factors and initial claims: DI-AR, lag with Beta weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
January	1.565	1.460	0.872	1.080	1.103	0.986	0.992	0.940	0.904	0.930	1.312	0.872
February	1.061	0.743	0.992	1.224	1.401	1.259	0.941	0.971	1.018	0.764	0.974	1.147
March	2.535	1.095	0.693	1.429	1.601	1.128	1.519	0.995	1.096	0.790^{*}	1.354	0.883
April	1.218	3.328	0.857^{*}	1.398	1.172	0.922	1.153	0.908	1.035	1.053	1.219	0.786^{*}
May	2.339	2.182	1.744	1.303	0.945	1.247	1.231	1.081	1.230	1.028	1.515	1.049
June	2.915	1.408	1.308	1.348	1.144	1.108	1.096	1.123	1.110	0.916	0.939	1.044
July	1.307	1.259	2.012	1.312	1.332	0.697	1.198	1.183	1.010	1.120	1.188	1.814
August	3.671	0.646	1.586	1.556	1.378	2.155	1.200	0.642	1.126	0.947	0.946	0.895
September	3.477	1.225	1.342	1.238	1.288	0.723	1.285	0.845	1.328	1.027	1.019	0.974
October	1.884	1.831	1.873	1.919	1.927	1.056	1.072	1.297	0.875^{**}	0.806	1.335	1.005
November	1.920	1.742	2.294	1.672	2.091	1.607	1.498	1.143	1.386	1.078	0.721	1.250
December	1.155	1.106	0.884	0.920	1.120	0.647^{*}	0.992	0.717^{**}	1.109	1.137	0.770^{**}	1.124

Notes: The stars denote statistical significance at 10%(*), 5%(**) and 1%(***) level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

Table 31: Factors and initial claims: DI-AR with Beta weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
January	1.462	1.290	1.011	1.241	1.048	1.141	0.946	0.930	0.999	1.135	1.208	1.042
February	0.707	0.607	1.083	1.241	1.380	1.113	1.031	0.997	0.975	1.007	1.160	1.347
March	1.345	1.333	0.457	1.423	1.402	0.992	1.145	1.027	1.081	0.803^{*}	1.066	1.062
April	0.602^{*}	1.754	0.951	1.316	1.124	1.031	1.152	1.042	1.052	1.019	0.909	0.837^{**}
May	1.128	1.892	2.059	1.146	0.918	1.132	1.072	0.983	1.071	1.076	1.029	0.953
June	1.822	1.084	1.344	1.455	1.020	1.255	1.255	1.061	1.128	1.002	0.995	0.990
July	2.606	1.414	1.654	1.294	1.231	0.747	1.068	1.071	0.950	1.118	1.037	1.177
August	1.850	1.295	1.028	1.409	1.079	1.400	0.855	0.647	1.161	0.884	0.985	0.933
September	2.147	0.721	0.869	0.997	0.940	0.760	1.344	0.747	1.059	1.081	0.888	1.021
October	1.597	1.660	1.367	1.614	1.195	0.893	0.960	1.153	0.898^{*}	0.841^{**}	1.037	0.949
November	1.028	1.524	1.411	1.248	1.204	1.245	1.122	0.877	1.234	0.964	0.692	1.125
December	1.225	1.552	0.739	1.119	1.098	0.722^{*}	0.920	0.839	0.755	1.118	0.797^{***}	1.061

Notes: The stars denote statistical significance at 10%(*), 5%(**) and 1%(***) level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

Table 32: Factors and initial claims: DI with Beta weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
January	1.130	1.381	1.287	1.251	1.151	1.166	0.938^{*}	1.037	1.030	1.038	1.234	1.131
February	0.729	0.513	1.028	1.230	1.140	0.928	1.037	0.946	0.981	1.077	1.053	1.272
March	1.397	0.943	0.610	1.149	1.309	0.957	0.947	0.995	1.040	0.910	0.927	0.959
April	0.572^{*}	1.544	0.864^{*}	0.879	1.284	1.027	1.116	1.005	1.069	1.005	0.899	0.711^{*}
May	1.321	1.442	2.025	0.876	1.073	1.120	1.109	0.929	0.968	1.046	1.100	0.983
June	1.552	1.108	1.085	1.545	1.020	0.986	1.247	1.114	1.079	1.014	1.028	1.006
July	2.273	1.059	1.233	1.394	1.782	0.683	0.986	1.169	1.074	1.093	1.049	1.096
August	2.734	0.763	1.042	1.464	1.177	1.729	0.770	0.649	1.023	0.981	0.958	0.903
September	2.604	1.366	0.667	0.847	1.184	0.714	1.574	0.679	0.898	1.121	0.990	1.011
October	1.830	1.867	1.770	1.249	1.185	1.071	0.937	1.114	0.862	0.879	1.032	0.953
November	0.978	1.396	1.191	1.355	0.990	1.081	1.185	0.940	1.288	1.009	0.793	1.031
December	1.155	1.633	0.865	0.992	1.104	0.756	0.983	0.862	0.798	1.197	0.896^{**}	1.137

Notes: The stars denote statistical significance at 10%(*), 5%(**) and 1%(***) level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

Table 33: Factors and initial claims: DI-AR, lag with Nealmon weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
January	1.671	1.395	0.898	1.040	1.031	1.078	1.047	0.940	1.106	0.835	1.366	0.872
February	0.791	0.725	1.065	1.313	1.479	1.181	0.963	0.934	1.055	0.803	0.986	1.322
March	2.278	1.065	0.952	1.601	1.675	0.976	1.455	1.005	1.012	0.837	1.457	0.863^{*}
April	1.287	3.168	0.668	1.243	1.291	1.020	1.136	0.906	1.045	1.046	1.299	0.937
May	2.780	1.919	2.064	1.280	1.075	1.379	1.331	1.025	1.186	1.032	1.328	0.902
June	2.246	1.868	1.238	1.236	1.054	1.151	1.117	1.182	1.047	0.883	0.938	0.953
July	2.500	0.900	1.826	1.186	1.084	0.784	1.249	1.233	1.055	1.166	1.216	1.613
August	4.236	0.775	1.549	1.628	1.488	2.048	1.175	0.649	0.978	0.958	1.105	0.993
September	4.492	1.866	1.433	1.504	1.286	0.798	1.508	0.867	1.341	0.991	1.103	1.162
October	2.775	2.399	1.684	1.964	2.113	0.952	1.119	1.611	0.904^{*}	0.915	1.257	1.117
November	1.724	1.707	1.833	1.782	1.985	1.617	1.575	1.069	1.521	1.188	0.615	1.376
December	1.218	1.269	1.000	0.935	1.210	0.780	1.011	0.761^{**}	1.089	1.380	0.764^{**}	1.079

Notes: The stars denote statistical significance at $10\%(^{*})$, $5\%(^{**})$ and $1\%(^{***})$ level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

Table 34: Factors and initial claims: DI-AR with Nealmon weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
January	1.232	1.349	0.916	1.250	1.062	1.114	1.009	0.884	0.999	1.163	1.159	1.054
February	0.612	0.659	1.076	1.201	1.381	1.174	1.050	0.948	0.977	0.966	1.186	1.326
March	1.589	1.033	0.498	1.400	1.444	0.984	1.183	1.038	1.078	0.875	1.023	1.009
April	0.509^{**}	1.629	1.006	1.307	1.074	0.949	1.173	1.033	1.051	1.061	0.909	0.861^{**}
May	1.368	2.090	1.887	0.976	0.841	1.159	1.038	0.977	1.077	1.071	1.027	0.958
June	1.822	1.408	1.408	1.678	1.058	1.165	1.264	1.094	1.130	0.995	0.990	0.954
July	3.169	1.420	1.521	1.410	1.184	0.950	0.992	1.107	0.909	1.147	1.006	1.149
August	1.947	1.261	1.028	1.354	0.978	1.562	0.824	0.613	1.143	0.847	0.987	0.974
September	1.629	0.707	0.769	0.944	0.930	0.786	1.266	0.734	1.038	1.031	0.874	1.012
October	1.658	1.716	1.408	1.501	1.182	0.918	0.984	1.079	0.965	0.992	1.020	0.943
November	1.000	1.551	1.344	1.212	1.112	1.241	1.169	0.886	1.258	0.937^{**}	0.714	1.158
December	1.258	1.721	0.739	1.024	1.109	0.747	0.889	0.817	0.822	1.105	0.786^{***}	0.994

Notes: The stars denote statistical significance at $10\%(^{*})$, $5\%(^{**})$ and $1\%(^{***})$ level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

Table 35: Factors and initial claims: DI with Nealmon weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
January	1.174	1.425	1.200	1.229	1.168	1.134	0.940^{**}	1.021	1.038	1.113	1.183	1.035
February	0.685	0.679	1.072	1.195	1.079	0.985	1.051	0.948	0.976	1.064	1.087	1.187
March	1.464	0.907	0.688	1.122	1.352	0.978	0.959	0.970	1.006	0.944	0.920	0.908
April	0.572^{*}	1.519	0.850	0.801	1.271	0.954	1.121	1.011	1.069	1.011	0.895	0.721^{*}
May	1.335	1.487	1.860	0.839	1.133	1.103	1.110	0.924	0.968	1.042	1.026	0.918^{*}
June	1.430	1.124	1.185	1.603	1.036	0.988	1.258	1.113	1.080	1.007	1.012	1.006
July	2.466	1.042	1.336	1.617	1.750	0.742	0.925	1.167	1.073	1.090	1.023	1.079
August	2.596	0.829	1.014	1.589	1.165	1.755	0.711	0.627	1.025	0.974	0.923	0.900
September	2.638	1.366	0.603^{*}	0.892	1.143	0.743	1.508	0.722	0.888	1.125	0.962	1.010
October	1.597	1.602	1.732	1.191	1.231	1.081	0.968	1.100	0.844	0.900	1.053	0.983
November	1.049	1.426	1.174	1.343	0.973	1.079	1.173	0.901	1.280	1.009	0.820	1.024
December	1.147	1.394	0.868	0.932	1.081	0.749	0.961	0.868	0.848	1.155	0.877^{***}	1.018

Notes: The stars denote statistical significance at $10\%(^{*})$, $5\%(^{**})$ and $1\%(^{***})$ level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

A.6.6. Factors, Google Trends index and initial claims

Table 36: Factors, Google Trends index and initial claims: DI-AR, lag with Beta weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
January	1.732	1.890	1.206	0.824	1.375	1.124	1.035	0.913*	0.976	1.024	1.033	0.969
February	1.118	0.795	0.863	1.123	1.175	1.109	0.959	0.974	1.018	0.779*	1.035	1.218
March	2.000	0.966	0.919	1.211	1.542	1.084	1.409	1.017	1.169	0.949	1.192	0.890
April	1.491	4.114	0.925	1.432	1.106	0.850	1.151	0.948	1.039	1.031	1.119	0.803*
May	1.716	2.084	1.908	1.385	1.052	1.192	1.022	1.129	1.058	1.046	1.293	0.957
June	2.611	1.270	1.276	1.257	1.175	1.159	1.157	1.020	1.002	1.042	0.966	1.160
July	1.683	1.420	2.288	1.364	1.377	0.749	1.389	0.972	0.903	1.127	1.071	1.933
August	3.804	0.639	1.115	1.679	1.509	2.705	1.268	0.606	1.034	0.902	0.894*	0.836
September	4.384	1.092	1.483	1.219	1.430	0.728	1.439	0.886	1.326	0.981	0.939	0.967
October	3.178	2.013	2.225	1.852	1.657	1.105	1.212	1.221	0.964	0.946	1.212	0.881
November	2.148	2.728	3.222	1.888	1.836	1.505	1.727	0.967	1.296	1.016	0.774	1.171
December	1.366	1.106	0.919	0.881	1.109	0.692	0.905	0.732**	1.042	1.042	0.731**	1.174

Notes: The stars denote statistical significance at 10%(*), 5%(**) and 1%(***) level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

Table 37: Factors, Google Trends index and initial claims: DI-AR with Beta weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
January	1.462	1.290	1.078	1.230	1.034	1.235	0.953	0.931	1.090	1.172	1.222	1.055
February	0.729	0.725	1.007	1.171	1.298	1.099	1.052	0.993	0.990	1.135	1.154	1.377
March	1.633	1.265	0.521	1.346	1.432	1.042	1.086	1.050	1.041	0.784*	1.290	1.131
April	0.719	1.871	0.925	1.290	1.097	1.016	1.218	0.993	1.048	1.027	0.930	1.049
May	1.191	1.799	1.897	1.110	0.867	1.072	1.098	1.028	1.000	1.111	0.930	1.119
June	1.651	1.177	1.350	1.576	1.016	1.296	1.237	1.113	1.169	0.966	0.983	1.057
July	2.606	1.218	1.774	1.206	1.158	0.723	1.016	1.081	0.955	1.131	1.024	1.200
August	2.176	1.322	1.007	1.475	1.105	1.161	0.882	0.597	1.206	0.927	0.991	0.984
September	2.750	0.635*	0.858	0.923	1.055	0.732	1.408	0.737	1.052	1.074	0.839	0.993
October	1.703	2.175	1.573	1.644	1.231	0.871	1.005	1.122	0.889*	0.874**	1.010	0.909
November	1.195	1.682	1.404	1.229	1.153	1.242	1.140	0.767	1.250	0.973	0.672	1.113
December	1.378	2.046	0.585	1.510	1.171	0.721*	0.941	0.842	0.700	1.114	0.805***	1.071

Notes: The stars denote statistical significance at 10%(*), 5%(**) and 1%(***) level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

Table 38: Factors, Google Trends index and initial claims: DI with Beta weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	*h=11*	*h=12*
January	1.130	1.468	1.372	1.329	1.297	1.166	0.926**	1.035	0.969	0.981	1.249	1.160
February	0.637	0.574	1.044	1.283	1.083	0.962	1.035	0.947	0.978	1.089	1.089	1.259
March	1.633	0.931	0.610	1.292	1.393	1.001	0.973	0.964	1.002	0.926	0.973	0.991
April	0.754	1.532	0.898	0.869	1.326	1.023	1.136	1.035	1.061	1.001	0.919	0.733*
May	1.160	1.442	1.876	0.873	1.066	1.098	1.087	0.934	0.931	1.050	1.066	1.009
June	1.492	0.955	1.122	1.586	0.946	0.963	1.287	1.108	1.142	0.998	1.051	1.102
July	2.198	1.075	1.116	1.402	1.696	0.701	0.995	1.191	1.031	1.133	1.040	1.147
August	1.451	0.996	1.062	1.309	1.171	1.644	0.783	0.609	1.059	0.951	0.950	0.893
September	2.476	0.572*	0.539*	0.793	1.121	0.706*	1.581	0.705	0.943*	1.166	0.961	1.012
October	1.378	2.427	1.408	1.338	1.200	0.969	0.958	1.125	0.907*	0.885	1.032	0.931
November	1.115	1.340	1.438	1.240	1.053	1.082	1.158	0.779	1.344	1.040	0.821	1.044
December	1.197	1.836	0.686	1.197	1.070	0.744	0.961	0.857	0.787	1.197	0.896**	1.149

Notes: The stars denote statistical significance at 10%(*), 5%(**) and 1%(***) level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

Table 39: Factors, Google Trends index and initial claims: DI-AR, lag with Nealmon weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
January	1.948	1.541	0.966	0.935	1.219	1.151	1.051	0.931^{**}	1.090	1.093	1.401	1.035
February	0.661	0.795	1.107	1.142	1.460	1.289	0.973	0.987	1.093	0.798	1.034	1.341
March	2.459	1.043	0.968	1.288	1.551	0.927	1.398	0.992	1.003	1.133	1.020	0.914
April	1.444	4.053	0.631	1.208	1.027	0.840	1.281	0.939	1.110	1.042	1.221	1.040
May	2.844	1.974	2.354	1.294	1.078	1.334	0.979	1.009	1.040	1.057	1.287	1.023
June	1.638	1.499	1.191	1.078	1.124	1.036	1.182	1.057	1.063	1.018	0.896	0.928
July	2.062	0.947	2.051	1.244	1.324	0.705	1.179	1.213	0.952	1.180	1.054	1.504
August	3.832	0.769	1.404	1.581	1.498	2.971	1.317	0.669	1.030	0.897	1.010	0.896
September	1.945	1.813	1.498	1.681	1.337	0.607^{*}	1.350	0.804	1.306	1.013	1.051	1.057
October	2.864	2.552	1.361	1.866	1.584	1.099	1.223	1.524	0.918	0.942	1.143	0.963
November	1.502	2.123	1.707	1.554	2.122	1.523	1.412	0.909	1.261	1.310	0.581	1.356
December	2.480	1.944	0.791	1.086	1.346	0.792	1.003	0.800^{**}	0.988	1.337	0.636^{***}	1.089

Notes: The stars denote statistical significance at $10\%(^{*})$, $5\%(^{**})$ and $1\%(^{***})$ level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

Table 40: Factors, Google Trends index and initial claims: DI-AR with Nealmon weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
January	1.365	1.428	1.019	1.241	1.132	1.211	1.017	0.873^{*}	1.061	1.196	1.125	1.060
February	0.707	0.707	0.952	1.129	1.304	1.144	1.048	0.960	0.975	1.086	1.158	1.377
March	1.952	1.075	0.564	1.347	1.472	1.057	1.064	1.036	1.043	0.848	1.261	1.085
April	0.799	1.961	0.925	1.325	1.080	0.961	1.216	0.990	1.044	1.062	0.954	1.087
May	1.183	1.885	1.892	1.003	0.835	1.120	1.079	1.024	1.018	1.105	0.994	1.053
June	1.758	1.499	1.320	1.688	1.054	1.232	1.314	1.150	1.165	0.962	0.983	0.995
July	3.227	1.390	1.572	1.289	1.065	0.935	0.974	1.089	0.913	1.128	1.022	1.210
August	2.224	1.332	1.121	1.411	1.112	1.460	0.830	0.577	1.225	0.895	0.966	0.989
September	2.670	0.772	0.798	0.909	1.038	0.750	1.387	0.734	1.036	1.023	0.842	1.005
October	1.761	2.013	1.443	1.548	1.221	0.879^{*}	0.998	1.042	0.929	1.013	1.000	0.912
November	1.276	1.732	1.379	1.229	1.068	1.273	1.161	0.817	1.296	0.967	0.708	1.165
December	1.432	2.046	0.658	1.395	1.139	0.732^{*}	0.889	0.814	0.787	1.085	0.733^{***}	0.995

Notes: The stars denote statistical significance at $10\%(^{*})$, $5\%(^{**})$ and $1\%(^{***})$ level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.

Table 41: Factors, Google Trends index and initial claims: DI with Nealmon weights.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
January	1.160	1.584	1.314	1.256	1.395	1.161	0.950^{**}	1.019	1.014	1.049	1.205	1.118
February	0.661	0.628	0.987	1.229	1.035	1.034	1.034	0.934	0.980	1.056	1.092	1.243
March	1.589	0.869	0.688	1.247	1.435	1.015	0.974	0.972	0.997	0.985	0.968	0.936
April	0.754	1.519	0.857	0.866	1.340	0.966	1.176	1.036	1.065	1.008	0.938	0.774^{*}
May	1.206	1.433	1.817	0.846	1.040	1.098	1.079	0.928	0.933	1.055	1.011	0.939
June	1.414	0.927	1.191	1.723	0.967	1.018	1.286	1.094	1.137	0.991	1.041	1.069
July	2.398	0.956	1.129	1.498	1.675	0.770^{*}	0.901^{**}	1.181	1.030	1.126	1.057	1.166
August	1.469	1.091	1.108	1.327	1.045	1.753	0.720	0.600	1.055	0.915	0.937	0.898
September	2.377	0.635^{*}	0.531^{*}	0.816	1.125	0.721^{*}	1.562	0.738	0.901^{*}	1.169	0.925	1.046
October	1.517	2.471	1.414	1.267	1.263	0.992	0.958	1.108	0.887^{*}	0.913	1.048	0.943
November	1.134	1.378	1.386	1.198	1.014	1.134	1.146	0.801	1.318	1.019	0.850	1.046
December	1.258	1.748	0.731	1.159	1.065	0.738	0.961	0.867	0.776	1.171	0.907^{**}	1.079

Notes: The stars denote statistical significance at $10\%(^{*})$, $5\%(^{**})$ and $1\%(^{***})$ level of the Diebold and Mariano (1995) test for the null of equal forecast accuracy between the benchmark and the competitor model when the benchmark is beaten.